



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS HUMANAS
DEPARTAMENTO DE GEOGRAFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA

USO DE SÉRIES TEMPORAIS DO SENSOR MODIS PARA IDENTIFICAR DIFERENTES CULTURAS AGRÍCOLAS

MIRIAM RODRIGUES DA SILVA

TESE DE DOUTORADO

Brasília - DF
Abril de 2018

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS HUMANAS
DEPARTAMENTO DE GEOGRAFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA

**USO DE SÉRIES TEMPORAIS DO SENSOR MODIS PARA
IDENTIFICAR DIFERENTES CULTURAS AGRÍCOLAS**

MIRIAM RODRIGUES DA SILVA

Orientador:
Prof. Dr. Osmar Abílio de Carvalho Júnior

TESE DE DOUTORADO

Brasília - DF
Abril de 2018

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS HUMANAS
DEPARTAMENTO DE GEOGRAFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA

USO DE SÉRIES TEMPORAIS DO SENSOR MODIS PARA IDENTIFICAR DIFERENTES CULTURAS AGRÍCOLAS

MIRIAM RODRIGUES DA SILVA

Tese de doutorado submetida ao Departamento de Geografia da Universidade de Brasília, como parte dos requisitos necessários à obtenção do Grau de Doutor em Geografia, área de concentração Gestão Ambiental e Territorial, opção Acadêmica.

Aprovado por:

Prof. Dr. Osmar Abílio de Carvalho Júnior (GEA/UnB)
(Orientador)

Prof. Dr. Luciano Shozo Shiratsuchi (Louisiana State University)
(Examinador Externo)

Prof. Dr. Raúl Sanchez Vicens (UFF)
(Examinador Externo)

Dr. André Luiz Farias de Souza (SIPAM)
(Examinador Externo)

Prof. Dr. Roberto Arnaldo Trancoso Gomes (GEA/UnB)
(Suplente)

Brasília - DF, 02 de abril de 2018.

FICHA CATALOGRÁFICA

SILVA, MIRIAM RODRIGUES DA

Uso de séries temporais do sensor MODIS para identificar diferentes culturas agrícolas. 89p. (UnB-IH-GEA, Doutora, Gestão Ambiental e Territorial, 2017).

Tese de Doutorado – Universidade de Brasília. Departamento de Geografia.

- | | |
|--------------------------------|--------------------------------|
| 1. Análise de séries temporais | 2. Redes neurais |
| 3. SVM | 4. Vizinho mais próximo (K-NN) |
| I. UnB-IH-GEA | II. Título (séire) |

REFERENCIA BIBLIOGRÁFICA

SILVA, M. R. Uso de séries temporais do sensor MODIS para identificar diferentes culturas agrícolas. 2018. 89p. Tese (Doutorado em Geografia) – Universidade de Brasília, Brasília, 2018.

CESSÃO DE DIREITOS

É concedida à Universidade de Brasília permissão para reproduzir cópias desta tese e emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta tese de doutorado pode ser reproduzida sem a autorização por escrito do autor.

Miriam Rodrigues da Silva

DEDICATÓRIA

Dedico este trabalho aos meus pais
Jorge e Maria, meus irmãos e
sobrinhos

AGRADECIMENTOS

Agradeço a Deus, pela oportunidade concedida, por ter me dado força e coragem para enfrentar e vencer mais esse desafio e me acompanhar em todos os momentos.

Aos meus pais Jorge e Maria pelo carinho e incentivo nos momentos difíceis. Aos meus irmãos Elias, Matusalém, Samoel, Daniel, David, Paulinho e Raquel por me apoiarem e torcerem sempre pelas minhas escolhas. Aos meus sobrinhos Verônica, Marlon, Paulinha, Samara, Thiago, Paulo Filho, Ana Clara, Débora, Giovanna, Gabriel, Danielle, João Pedro, Ana Clara, Geovana, Dafinny, Rafaella, Diogo, Isadora, Maria Eduarda e Júlia por todo amor a mim oferecido. Amo muito cada um de vocês, família!

Ao meu orientador Osmar Abílio Carvalho Júnior, pela confiança, paciência, dedicação, sugestões, amizade e ensinamentos transmitidos ao longo do curso.

Aos membros da banca professor Dr. Luciano Shozo Shiratsuchi (Louisiana State University), professor Dr. Raul Sanches Vicens (UFF) e Dr. André Luiz Farias de Souza (SIPAM) pelas contribuições para o aperfeiçoamento deste trabalho.

A APROSOJA (Associação dos Produtores de Soja e Milho de Mato Grosso), CONAB (Companhia Nacional de Abastecimento) e EMBRAPA Agrossilvipastoril por fornecerem os pontos GPS coletados em campo. Aos amigos Sandro Nunes de Oliveira e Nickolas Castro Santana pela realização do trabalho de campo. Ao pesquisador Luciano Shozo Shiratsuchi da EMBRAPA Agrossilvipastoril, ao Engenheiro Agrônomo Gabriel Ponce Alves da empresa Solo Planta e a todos os técnicos da CONAB pelo auxílio e logística no campo.

Um agradecimento em especial aos amigos Elisson Castro, Luis Antônio e Sandra Vanessa pelo carinho, paciência, por me acalmarem nos momentos de angústia e pelos momentos de descontração. E a amiga Aldenora Perrone que mesmo distante sempre torceu por mim.

Aos amigos do LSIE (Laboratório de Sistemas de Informações Espaciais) pelo companheirismo ao longo dessa jornada, pela troca de experiências, conversas nos momentos de incertezas e pela companhia durante a noite no laboratório: Verônica, Cristiane, Sandro, Nickolas, Marcus Fábio e Natanael. E pelo bom convívio dos colegas Potira, Jorge, Nathalia, Saiaka, Fernando, Anesmar, Marcos e Núbia.

Aos professores Renato Guimarães e Roberto Gomes, pelo suporte, convivência e incentivos dados neste período.

Ao Programa de Pós-graduação em Geografia da Universidade de Brasília, a todos os professores e aos funcionários da Secretaria de Pós-graduação.

Agradeço ao INMET, na pessoa do Dr. Antônio Divino Moura, pelo consentimento para fazer o curso e incentivo desta pesquisa. A amiga Camila Lopes, pela convivência, amizade e paciência ao longo desses anos e ao Divino Figueiredo pelas contribuições, troca de experiência e conversas.

Aos colegas da HEX Tecnologias Geoespaciais pela torcida: Thais Hagale, Thaise Rodrigues, Leonice, Pedro, Altamiro e William.

A todos, que torceram e contribuíram de alguma maneira para a realização desta tese, muito obrigada.

"Para se ter sucesso, é necessário amar de verdade o que se faz. Caso contrário, levando em conta apenas o lado racional, você simplesmente desiste. É o que acontece com a maioria das pessoas."

Steve Jobs

SUMÁRIO

LISTA DE FIGURAS	xi
LISTA DE TABELAS	xiii
LISTA DE ABREVIATURAS E SIGLAS	xv
RESUMO	xvii
ABSTRACT	xix
CAPÍTULO 1 - APRESENTAÇÃO GERAL	1
1.1. INTRODUÇÃO	1
1.2. ÍNDICE DE VEGETAÇÃO.....	4
1.3. CARACTERIZAÇÃO E APLICAÇÃO DO SENSOR MODIS	4
1.4. FENOLOGIA DA VEGETAÇÃO.....	8
1.5. ESTIMATIVA DE ESTÁGIOS FENOLÓGICOS DE CULTURAS AGRÍCOLAS	9
1.6. ANÁLISE DE SÉRIES TEMPORAIS.....	10
1.7. SUAVIZAÇÃO DAS SÉRIES TEMPORAIS	11
1.8 – MÉTODOS DE CLASSIFICAÇÃO	12
1.9. ANÁLISE DE ACURÁCIA.....	14
1.10. SISTEMAS DE MONITORAMENTO GLOBAL DA AGRICULTURA	16
1.11. AGRICULTURA NAS ÁREAS DE ESTUDO	16
1.12. ORGANIZAÇÃO DA TESE	23
REFERENCIAS BIBLIOGRÁFICAS	24
CAPÍTULO 2 – COMPARING BACK-PROPAGATION NEURAL NETWORKS AND SUPPORT VECTOR MACHINE FOR CROPPING SYSTEM MAPPING USING MODIS NDVI TIME SERIES IN MATO GROSSO STATE, BRAZIL	33
2.1. INTRODUCTION	33
2.2. MATERIALS AND METHODS	35
2.2.1. Study Area.....	35
2.2.2. MODIS Data	36

2.2.3. NDVI Time Series Noise Reduction.....	37
2.2.4. MODIS Time-Series Classification.....	37
2.2.5. Accuracy Analysis.....	38
2.3. RESULTS.....	39
2.3.1. Temporal signatures	39
2.3.2. Results of time-series classifications using ANN	40
2.3.3. Results of time-series classifications using Support Vector Machine.....	45
2.4. DISCUSSION.....	49
2.5. CONCLUSIONS	51
Acknowledgments:	51
REFERENCES	52
CAPÍTULO 3 – WHEAT PLANTED AREA DETECTION FROM THE MODIS NDVI TIME SERIES CLASSIFICATION USING THE NEAREST NEIGHBOR METHOD CALCULATED BY THE EUCLIDEAN DISTANCE AND COSINE SIMILARITY MEASURES.....	
3.1. INTRODUCTION	58
3.2. STUDY AREA	60
3.3. MATERIALS AND METHODS	61
3.3.1. MODIS data	61
3.3.2. Elaboration and filtering of the NDVI temporal cube.....	61
3.3.3. Selection and classification of temporal signatures.....	62
3.3.4. Accuracy analysis.....	63
3.4. RESULTS.....	63
3.4.1. Analysis of temporal signatures of wheat cultivation	63
3.4.2. Results of the ratings of the time series.....	64
3.5. DISCUSSION.....	66
3.6. CONCLUSION	67
REFERENCES	68

CAPÍTULO 4 – MAPEAMENTO DE ÁREAS COM CULTIVOS DE ARROZ NO ESTADO DE SANTA CATARINA, ANO SAFRA 2014/2015, UTILIZANDO A TÉCNICA DO VIZINHO MAIS PRÓXIMO (K-NN), NA SÉRIE TEMPORAL NDVI MODIS	74
4.1. INTRODUÇÃO.....	74
4.2. ÁREA DE ESTUDO	76
4.3.MATERIAL E METODOLOGIA	77
4.3.1. Dados MODIS.....	77
4.3.2. Classificação Baseada no Método do Vizinho Mais Próximo	78
4.3.3. Análise da Acurácia	79
4.4.RESULTADOS	79
4.5.DISSCUSSÃO.....	83
4.6.CONCLUSÃO.....	84
AGRADECIMENTOS	84
REFERÊNCIAS BIBLIOGRÁFICAS	84
CAPÍTULO 5– CONSIDERAÇÕES FINAIS	88

LISTA DE FIGURAS

Figura 1. 1- Estimativa da produção de grãos no Brasil ano 2015.....	2
Figura 1. 2- Localização da três áreas de estudo	17
Figura 1. 3- Estimativa da produção de grãos no Estado do Mato Grosso	19
Figura 1. 4- Estimativa da produção dos principais grãos distribuídos por mesorregiões	20
Figura 1. 5- Estimativa da produção dos principais grãos distribuídos na área de estudo	20
Figura 1. 6- Estimativa da produção de grãos no Estado do Rio Grande do Sul	21
Figura 1. 7- Estimativa da produção dos principais grãos distribuídos por mesorregiões	21
Figura 1. 8- Estimativa da produção de grãos no Estado de Santa Catarina	22
Figura 1. 9- Estimativa da produção dos principais grãos distribuídos por mesorregiões	22
Figura 1. 10- Produção de arroz na área de estudo.....	22
Figura 1. 11- Estrutura da tese.....	23
Figure 2.1. Location of the study area.....	36
Figure 2.2. Temporal signatures of land use and crop classes: (a), (b), (c) and (d).	40
Figure 2.3. Histogram of percentage of kappa coefficients considering the different ANN models described in Table 2.	41
Figure 2.4. Graphs of the following average behaviors: (a) training rate, (b) training momentum, and (c) training threshold contribution.....	43
Figure 2.5. The four best-fit models using back-propagation neural networks considering value variations for training threshold contribution (v), training rate (η), and training momentum (μ): (a) first model ($v=0.7$, $\eta=0.2$, $\mu=0.4$); (b) second model ($v=0.6$, $\eta=0.2$, $\mu=0.4$); (c) third model ($v=0.5$, $\eta=0.2$, $\mu=0.4$); and (d) fourth model ($v=0.9$, $\eta=0.2$, $\mu=0.05$).....	44
Figure 2.6 Classified images using Support Vector Machine using the following kernel types: (a) linear; (b) polynomial of degree 2; (c) polynomial of degree 3; (d) polynomial of degree 4; (e) polynomial of degree 5; (f) polynomial of degree 6; (g) RBF; and (h) sigmoid.	46
Figure 3. 1. Location map of the study area.....	60
Figure 3. 2. Filtering of the MODIS NDVI time series by the Savitzky-Golay method.....	62
Figure 3. 3. MODIS NDVI temporal signatures for wheat crop in the Northwest Rio Grande do Sul.....	64
Figure 3. 4. Kappa index curve for the definition of threshold values for the two metrics of the nearest neighbor: (a) Euclidean distance and (b) cosine similarity.	64

Figure 3. 5. Wheat planting map by using the threshold value in the nearest neighbor metric images: (a) Euclidian distance, (b) cosine similarity.....	66
Figura 4. 1. Produção brasileira de arroz em área plantada por hectare. Fonte: IBGE, 2016.	75
Figura 4. 2. Localização da área de estudo.....	76
Figura 4. 3. Curva da série temporal original e com a aplicação do filtro Savitzky-Golay.	78
Figura 4. 4. Variação temporal dos perfis de NDVI para cada grupo de arroz.	80
Figura 4. 5. Valor limite para a detecção da máscara de arroz a partir da imagem da métrica do vizinho mais próximo: (a) distância Euclidiana, e (b) semelhança de cossenos.	81
Figura 4. 6. Área de arroz classificada com: a) Distância Euclidiana e b) semelhança de cossenos.	82

LISTA DE TABELAS

Tabela 1. 1- Área e Produção de Grãos por Região — 1973/99	18
Table 2. 1 - Data layout for McNemar test between two classification results.....	39
Table 2. 2. Kappa coefficients of the classified images from different ANN models using logistic function. In the tests, the following constant parameters were used: simple hidden layer, training RMS exit criteria of 0.1, and number of maximum iterations of 5,000. The variable parameters were the training rate (η), training momentum (μ), and training threshold contribution (ν). The best-fit model is marked in red.	42
Table 2. 3. Accuracy assessment (Overall accuracy, Kappa coefficient and confusion matrix) for the four best ANN classifications considering value variations for training threshold contribution (ν), training rate (η), and training momentum (μ). All models used a simple hidden layer, training RMS exit criteria of 0.1, and number of maximum iterations of 5,000. The classifications considered seven classes (ASCS – Annual Single Cropping System, S/Ct – Soybean/Cotton, S/Cr – Soybean/Corn, CP– Center Pivot, P – Pasture, S – Savanna, and FF – Forest Formation.	43
Table 2. 4. Commission and omission errors for the four best ANN classifications and classes (ASCS – Annual Single Cropping System, S/Ct – Soybean/Cotton, S/Cr – Soybean/Corn, CP– Center Pivot, P – Pasture, S – Savanna, FF – Forest Formation).	45
Table 2. 5. McNemar’s test for the four best ANN classifications considering changes to the parameters: training threshold contribution (ν), training rate (η), and training momentum (μ). The values marked in red have calculated χ^2 values higher than the critical value at a significance level of 5% (3.841).	45
Table 2. 6. Accuracy assessment (Overall accuracy - OA, Kappa coefficient – K, and confusion matrix) for the classifications based on Support Vector Machine (SVM) using kernel functions (linear, polynomial, RBF, and sigmoid). The classifications considered seven classes (ASCS – Annual Single Cropping System, S/Ct – Soybean/Cotton, S/Cr – Soybean/Corn, CP– Center Pivot, P – Pasture, S – Savanna, FF – Forest Formation).	47
Table 2. 7. Commission and omission errors for the different SVM methods and classes (ASCS – Annual Single Cropping System, S/Ct – Soybean/Cotton, S/Cr – Soybean/Corn, CP– Center Pivot, P – Pasture, S – Savanna, FF – Forest Formation).	48
Table 2. 8. McNemar’s test among the different SVM models. The values marked in red have calculated χ^2 values higher than the critical value at a significance level of 5% (3.841).	48

Table 2. 9. McNemar's test between the best ANN model ($v=0.7$, $\eta=0.2$, $\mu=0.4$) and SVM models.....	49
Table 3.1. Confusion Matrix of the classifications based on the nearest neighbor metric images: Euclidean distance and cosine similarity. Where "EO" is the error of omission and "EC" is the error of commission.....	65
Table 3.2. McNemar's test between the two classifications by the nearest neighbor using Euclidean distance and the cosine similarity.	65
Tabela 4.1. Calendário de plantio (P) e colheita (C) trigo (Fonte: CONAB, 2016).....	77
Tabela 4.2. Área cultivada com arroz em hectares estimada com as duas metodologias baseada nas métricas do vizinho mais próximo: distância euclidiana (DE) e semelhança de cossenos (SC).....	82
Tabela 4.3. Matriz de confusão em percentagem da classificação baseada nas métricas de distância e similaridade do vizinho mais próximo.	82
Tabela 4.4. Teste de McNemar's entre as duas classificações do vizinho mais próximo usando as métricas de distância Euclidiana e semelhança de cossenos	83

LISTA DE ABREVIATURAS E SIGLAS

ANN - Artificial Neural Network

APROSOJA - Associação dos Produtores de Soja e Milho do Estado do Mato Grosso

AVHRR – Advanced Very High Resolution Radiometer

CAS – Crop Watch System

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

CONAB- Companhia Nacional de Abastecimento

CPPI - índice de proporção da fenologia da cultura

EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária

ENVI - Environment for the Visualization of Images

EPAGRI - Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina

ETM - Enhanced Thematic Mapper

EVI - *Enhanced Vegetation Index*

FAO - Food and Agriculture Organization

GEOGLAM - Group on Earth Observations Global Agricultural Monitoring

GIEWS - Global Information and Early Warning System

GLAM - Global Agriculture Monitoring

GPP - produtividade primária bruta

HDF - Hierarchy Data Format

IBGE – Instituto Brasileiro de Geografia e Estatística

ICA- Análise de Componentes Independentes

IV- Índice de Vegetação

K-NN –K-Nearest Neighbor

LANDSAT - Land Remote Sensing Satellite

LSIE - Laboratório de Sistemas de Informações Espaciais

LSWI – Land Surface Water Index

MARS - Monitoring Agriculture by Remote Sensing

MD - Mínima Distância

MODIS - *Moderate Resolution Imaging Spectroradiometer*

MRT - MODIS Reprojection Tools

NASA - National Aeronautics and Space Administration

NASS - National Agricultural Statistics Service

NDVI – *Normalized Difference Vegetation Index*

NIR - Near Infrared

NOAA - National Oceanic and Atmospheric Administration

RBF - radial basis function

RMS – Root–Mean Squared

SAM - *Spectral Angle Mapper*

SG - Savitzky and Golay

SIG - Sistemas de Informações Geográficas

SOSBAI - Sociedade Sul-Brasileira de Arroz Irrigado

SVM- Support Vector Machines

TM - Thematic Mapper

UnB - Universidade de Brasília

USGS - United States Geological Sur

WDRVI - Wide Dynamic Range Vegetation Index

WGS84 - World Geodetic System 84

RESUMO

A presente pesquisa objetiva identificar culturas de grãos a partir de séries temporais NDVI MODIS. As culturas agrícolas e regiões analisadas foram: (a) soja, milho e algodão no Estado do Mato Grosso na safra de 2013/2014; (b) trigo no Estado do Rio Grande do Sul; (c) e cultura do arroz no Estado de Santa Catarina. A tese está estruturada em 5 (cinco) capítulos, onde os capítulos de desenvolvimento (2, 3 e 4) foram escritos no formato de artigos científicos. No processamento digital de imagem todas as análises consideraram as seguintes etapas: (a) aquisição das imagens MODIS; (b) tratamento dos ruídos usando o filtro Savitzky-Golay; (c) classificação; e (d) análise de acurácia. A principal diferença metodológica foi a etapa de classificação que utilizou duas abordagens: (a) classificação contínua do terreno considerando as diferentes produções agrícolas (soja, milho e algodão) e os tipos de vegetação a partir de dois métodos de aprendizagem de máquina (Support Vector Machines e Redes Neurais de retro-propagação); e (b) detecção de uma única cultura de pequenos agricultores (arroz em Santa Catarina e trigo no Rio Grande do Sul) usando o método do vizinho mais próximo (caso específico do método K-NN). A primeira abordagem usando classificação contínua do terreno considerou as seguintes assinaturas temporais NDVI: formação florestal, cerrado, pastagem, sistema único de cultivo anual (soja, milho e algodão), sistema duplo de cultivo (soja/milho e soja/algodão) e pivô central (sistema triplo de cultivo). Na classificação foram testados 378 modelos de redes neurais com variações dos parâmetros de entrada e 8 modelos SVM usando diferentes funções Kernel. O índice Kappa mostrou que os melhores modelos da Rede Neural (0,77) e SVM (0,75) foram estatisticamente equivalentes pelo teste McNemar. A classificação baseada no vizinho mais próximo foi constituída de duas fases: (a) geração de imagens métricas (distância Euclidiana e similaridade do cosseno); e (b) definição do melhor valor de corte para caracterizar a máscara da cultura agrícola. Os resultados mostraram diferentes perfis temporais tanto no trigo como no arroz devido às variações do calendário agrícola da região. Nas duas classificações (trigo e arroz), os resultados usando as duas métricas foram estatisticamente equivalentes pelo teste McNemar. Na análise do trigo, a distância Euclidiana obteve um índice Kappa de 0,75 e a semelhança do cosseno um índice Kappa de 0,74. Na análise do arroz a distância Euclidiana obteve um índice Kappa de 0,73 e a semelhança do cosseno um índice Kappa de 0,72. As metodologias descritas demonstram uma grande potencial para o cálculo das áreas de produção agrícola, podendo auxiliar os órgãos federais para o planejamento regional e segurança alimentar.

Palavras chave: MODIS NDVI, análise de séries temporais, fenologia, redes neurais, SVM, vizinho mais próximo (K-NN), distância Euclidiana, similaridade do cosseno.

ABSTRACT

The present research aims to identify grain crops from NDVI MODIS time series. The agricultural crops and the analyzed regions were: (a) soybean, corn and cotton in Mato Grosso State at 2013/14 growing season; (b) wheat in the State of Rio Grande do Sul; (c) and rice in the State of Santa Catarina. The thesis is structured in 5 (five) chapters, where the development chapters (2, 3 and 4) were written in the format of scientific articles. In digital image processing, all analyzes considered the following steps: (a) acquisition of MODIS images; (b) noise treatment using the Savitzky-Golay filter; (c) classification; and (d) accuracy analysis. The main methodological difference was the classification stage that used two approaches: (a) continuous land classification considering the different agricultural production (soybean, corn and cotton) and vegetation types from two methods of machine learning (Support Vector Machines and Retro-propagation Neural Networks); and (b) detection of a single crop of small farmers (rice in the Santa Catarina and wheat in the Rio Grande do Sul) using the nearest neighbor method (specific case of the K-NN method). The first approach using continuous land classification considered the following NDVI temporal signatures: forest formation, cerrado, pasture, single annual cropping system (soybean, corn and cotton), double cropping system (soybean / corn and soybean / cotton) and pivot (triple cropping system). In the classification were tested 378 models of neural networks with different variations in input parameters and 8 SVM models using different Kernel functions. The Kappa index showed that the best models of the Neural Network (0.77) and SVM (0.75) were statistically equivalent by the McNemar test. The classification based on the nearest neighbor was constituted of two phases: (a) elaboration of metric images (Euclidean distance and similarity of the cosine); and (b) definition of the best threshold value to characterize the agricultural crop mask. The results showed different temporal profiles in both wheat and rice due to variations in the region's agricultural calendar. In both classifications (wheat and rice), the results using the two metrics were statistically equivalent by the McNemar test. In wheat analysis, the Euclidean distance obtained a Kappa index of 0.75 and the cosine similarity a Kappa index of 0.74. In rice analysis, the Euclidean distance obtained a Kappa index of 0.73 and the cosine similarity a Kappa index of 0.72. The methodologies showed a promising potential to determine the areas of crop production and could be very useful for federal agencies for regional planning and food security programs.

Keywords: MODIS NDVI, time series analysis, phenology, neural networks, SVM, nearest neighbor (K-NN), Euclidean distance, cosine similarity.

CAPÍTULO 1 - APRESENTAÇÃO GERAL

1.1. INTRODUÇÃO

A agricultura brasileira passou por um processo de transformação nos meados da década de 1960 até o final da década de 1980. Essas transformações foram possibilitadas pela modernização da agricultura, pela expansão do crédito rural subsidiado destinado a aquisição de insumos modernos, e pelo processo de substituição de importações, tanto na área mecânica quanto na área química de insumos agrícolas. Esse processo colaborou para o fortalecimento do setor de insumos modernos, que cresceu na direção da expansão da área plantada, das vantagens à produção doméstica de insumos, do aumento do volume de crédito seletivo e do incentivo das exportações agrícolas (CONCEIÇÃO & CONCEIÇÃO, 2014).

A importância da produção agrícola no Brasil é decisória na economia do país. Desde o ano 2000, o agronegócio brasileiro vem ganhando participação no mercado internacional de alimentos, graças ao aumento contínuo do volume exportado, e também nas exportações totais brasileiras, as quais já representam 40% (BARROS & ADAMI, 2013). A partir da segunda metade do século 20, o agronegócio brasileiro passou por diversas transformações, conduzido não apenas pela modernização tecnológica da agricultura, mas também por outros fatores relevantes, tais como a abertura da economia internacional e a globalização, impondo novas condições à competitividade (EMBRAPA, 2003). Com o processo de globalização da agricultura, aconteceram grandes transformações em todo o sistema de produção, solicitando mais competitividade em todos os segmentos relacionados. Assim, os processos produtivos da agricultura reuniram novos conceitos, métodos e tecnologias, ocasionando importantes mudanças.

O Centro-Oeste é a principal região produtora de grãos no país, destacando-se as culturas de milho, soja e algodão. Esta região é responsável por 40,3% da produção nacional, seguidos da região Sul com 38,1%. O Estado do Mato Grosso possui uma participação de 23,6% da produção, liderando como o maior Estado produtor brasileiro, seguido pelo Estado do Paraná com 18,5% e Rio Grande do Sul com 16,30% (**Figura 1.1**) (IBGE, 2015).

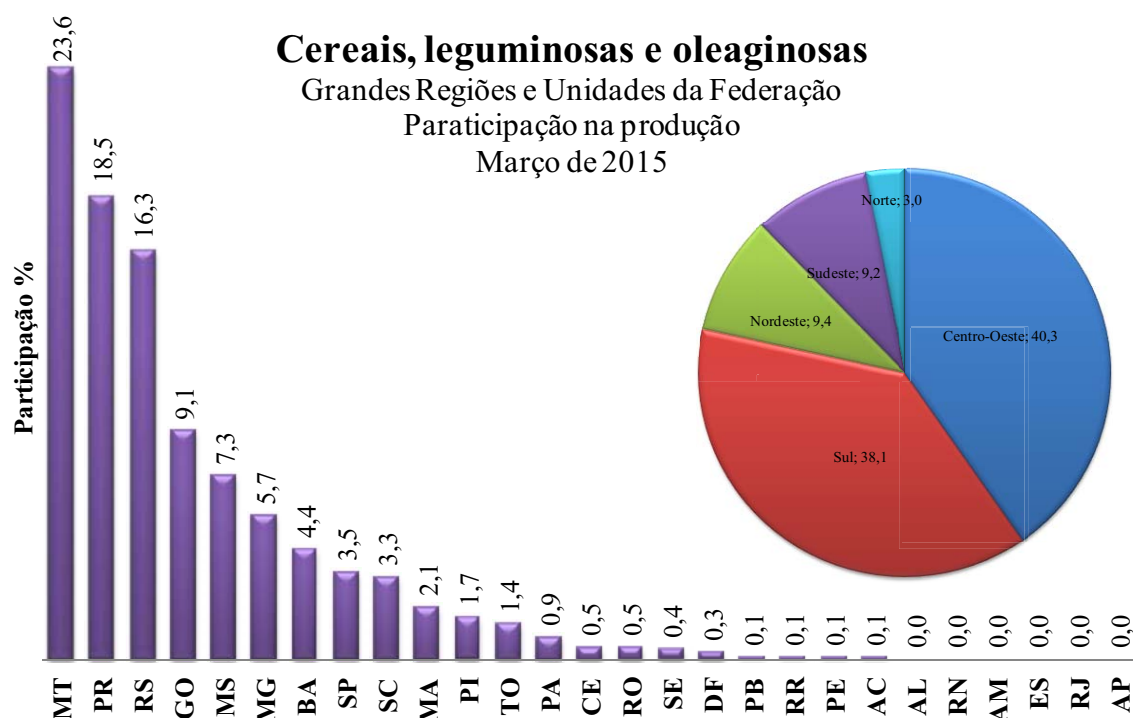


Figura 1. 1. Estimativa da produção de grãos no Brasil ano 2015.

Fonte: IBGE, 2015.

Para se estimar a produção agrícola com as principais lavouras de grãos no Brasil, são utilizados métodos baseados em informações obtidas juntos aos agricultores e entidades relacionadas ao setor agrícola como a CONAB (Companhia Nacional de Abastecimento) e o IBGE (Instituto Brasileiro de Geografia e Estatística). Contudo, a estratégia empregada para este tipo de cálculo exige muita mão de obra e alto custo, além de estar susceptível a erro humano. Entretanto, vários trabalhos têm demonstrado que imagens de satélites apresentam um grande potencial para mapear e estimar áreas agrícolas (AGUIAR et al., 2008; LOHMANN et al., 2009; ROSA et al., 2010; PAN et al., 2012; PENG et al., 2013). Diante disso, as geotecnologias tornaram-se uma importante ferramenta para o monitoramento e a previsão de safras de grandes áreas agrícolas, diminuindo os custos e o tempo de levantamento de informações (PIFFER et al., 2009; 2010; AMARAL et al., 2011; SANTOS et al., 2011; SOUSA et al., 2011; SOUZA et al., 2011; MARQUEZOLA & SANTOS, 2013). Uma das formas de monitoramento em escala global da superfície terrestre se dá por meio da utilização de sensores como o MODIS (*Moderate Resolution Imaging Spectroradiometer*) (SAKAMOTO et al., 2005; CARVALHO JÚNIOR et al., 2008; GOLFORD et al., 2008; KLERING et al., 2009; OZDOGAN, 2010; SANTANA et al., 2010; ARVOR et al., 2011; COUTO JÚNIOR et al., 2012).

Sendo assim, este trabalho tem o escopo de colaborar no desenvolvimento de metodologias operacionais para avaliação de culturas agrícolas a partir do uso de geotecnologias. Para isto, tentaremos responder três perguntas:

1 - A disponibilidade de imagens oriundas do produto MOD09Q1 do sensor MODIS facilita a identificação e mapeamento das culturas agrícolas no Brasil?

2 - O uso de parâmetros espectrais, como índices de vegetação e imagens de reflectância de superfície proveniente de imagens multitemporais do sensor MODIS, permitem a identificação e estimativa da área plantada de forma operacional?

3 - Com o NDVI do MODIS é possível identificar o comportamento espectral das diversas culturas agrícolas?

Devido à importância do agronegócio para a economia do país, ações governamentais para o controle da balança comercial dos produtos agrícolas, o apropriado abastecimento do mercado interno, demandam informações frequentes e confiáveis sobre a produção agrícola. Estas informações são indispensáveis para orientar financiamentos agrícolas e para que o agricultor possa tomar decisões acertadas no que se refere ao plantio e a comercialização de seus produtos. Porém, a grande extensão do território brasileiro, associado a dinâmica espaço-temporal da atividade agrícola, dificulta a aquisição de informações confiáveis sobre este setor da economia. Em compensação, a utilização do sensoriamento remoto e do SIG, tem a capacidade de auxiliar na obtenção dessas informações de forma rápida, precisa e com menor custo (RIZZI & RUDORFF, 2005).

Deste modo, trabalhos têm sido desenvolvidos objetivando o monitoramento agrícola e previsão de safras a partir do uso de geotecnologias e modelos de estimativas de produção, baseados em parâmetros espectrais e agrometeorológicos, conferindo credibilidade as estimativas das áreas plantadas.

Assim, o objetivo geral da tese é:

✓ Distinguir culturas específicas de grãos, utilizando as séries temporais do sensor MODIS e identificar onde estão localizadas estas culturas, por meio de algoritmos de classificação.

E os objetivos específicos são:

✓ Comparar os métodos SVM e Rede Neurais de retro-propagação aplicado a séries temporais MODIS, considerando as regiões produtoras de soja, milho e algodão na Mesorregião Norte do Estado do Mato Grosso, durante ano safra de 2013/2014.

✓ Distinguir a cultura do trigo ano safra 2013, na Mesorregião Noroeste do Rio Grande do Sul usando o método de classificação do vizinho mais próximo (K-NN) a partir de duas métricas: similaridade e distância euclidiana.

✓ Identificar a cultura do arroz, usando o método de classificação do vizinho mais próximo (K-NN) a partir de duas métricas: similaridade e distância euclidiana, ano safra 2014/2015, no Estado de Santa Catarina.

1.2. ÍNDICE DE VEGETAÇÃO

Os Índices de Vegetação (IVs) são alterações espectrais de duas ou mais bandas, aplicados para aumentar a contribuição da propriedade da vegetação e possibilitar comparações, inter-espaciais e temporais, confiáveis da atividade fotossintética terrestre (HUETE et al., 2002). São usados frequentemente para monitorar e caracterizar a cobertura vegetal da Terra, relacionando-se com parâmetros biofísicos da vegetação, como biomassa, índice de área foliar, porcentagem da cobertura verde e teor de clorofila (JIANG et al., 2008).

Dentro os índices de vegetação utilizados para o monitoramento da vegetação destacam-se o *Normalized Difference Vegetation Index* (NDVI) e *Enhanced Vegetation Index* (EVI) (HUETE et al., 2002). No geral, o NDVI é mais sensível à presença de pigmentos que participam dos processos fotossintéticos, em particular ao conteúdo de clorofila e pode ser aplicado em diferentes tipos de vegetação (GITELSON et al., 1996; HUETE et al., 1997). O NDVI decorre como uma medida de vegetação em que ele é suficiente estável para permitir comparações significativas de mudanças sazonais e inter-anuais de crescimento da vegetação. No entanto, uma peculiaridade inerente ao NDVI é a saturação assintótica, o que o torna pouco sensível à detecção de variações de biomassa verde em condições de elevados índices de área foliar (HUETE et al., 2002). O NDVI reduz algumas fontes externas de ruído como, por exemplo, sombras de nuvens (JUSTICE et al., 1998). O EVI foi desenvolvido para aprimorar o sinal da vegetação e minimizar os efeitos tanto do solo quanto das influências de sinal de fundo do dossel e da atmosfera (JUSTICE et al., 1998; HUETE et al., 1999).

1.3. CARACTERIZAÇÃO E APLICAÇÃO DO SENSOR MODIS

Durante a última década, o MODIS tornou-se uma ferramenta fundamental para o monitoramento da vegetação devido à sua alta resolução temporal, extensa abrangência e rápida disponibilidade de vários produtos.

As imagens oriundas do sensor MODIS a bordo dos satélites Terra e Aqua, dada sua resolução temporal de recobrimento, beneficia a aquisição de imagens livres de cobertura de nuvens e viabilizam o monitoramento da cultura com maior frequência de observação, ao

longo de seu ciclo vegetativo e produtivo. Dentre os produtos gerados a partir dos dados coletados por este sensor, os índices de vegetação NDVI e EVI contidos no produto MOD13Q1, são voltados especialmente para vegetação, fornecidos a cada 16 dias, georreferenciados e corrigidos para os efeitos atmosféricos, com resolução espacial de 250 x 250m (JUSTICE et al., 2002; VERMOTE et al., 2002).

O estudo com séries temporais do sensor MODIS vem sendo aplicado em áreas naturais (ZHANG et al., 2003; COUTO JÚNIOR et al., 2007; CARVALHO JÚNIOR et al., 2008a; SANTANA et al., 2010; BORGES et al., 2014) e em áreas agrícolas (D'ARCO et al., 2007; AGUIAR et al., 2008; GALFORD et al., 2008; LOHMANN et al., 2009; OZDOGAN, 2010; ROSA et al., 2010; COUTO JÚNIOR et al., 2012; PAN et al., 2012; PENG et al., 2013).

Zhang et al., (2003) estudaram processos da fenologia da vegetação para o nordeste dos Estados Unidos, a partir do sensor MODIS com o uso de séries temporais com os dados de índice de vegetação (IV) e identificaram o comportamento fenológico caracterizado por vários períodos de crescimento e senescência em um único ano. Segundo os autores este método tem potencial para o monitoramento da fenologia da vegetação quase em tempo real.

Couto Júnior et al., (2007) identificaram as assinaturas temporais das formações vegetais presentes na região do EcoMuseu do Cerrado e as formações florestais apresentaram maiores valores de NDVI na série temporal e a menor variação entre as estações enquanto que as formações savânicas e campestres mostraram uma menor diferença entre os valores de NDVI.

Em uma classificação de savana usando assinaturas temporais NDVI do sensor MODIS no Parque Nacional Chapada dos Veadeiros, Carvalho Júnior et al., (2008) mostraram que as séries NDVI são úteis na diferenciação dos diferentes tipos de vegetação.

Em um estudo sobre a modelagem de espectros temporais NDVI-MODIS, no período de 2000 a 2008, na bacia do Rio Paracatu, Santana et al., (2010) modelaram a distribuição de dados em seis diferentes alvos: área de plantio, Floresta de Galeria, Cerrado, pastagem, área urbana e Floresta Estacional Semidecidual e também analisaram as mudanças fisionômicas dos índices por meio dos modelos ondulatórios. Identificaram as fisionomias a serem delimitadas e simularam séries temporais futuras, pelo uso da terra, coberturas vegetais e o tamanho da área de cada fisionomia.

Borges & Sano (2014) utilizaram a série temporal do MODIS para demonstrar o potencial do sensor na discriminação de uso e cobertura vegetal presentes no Oeste da Bahia. Utilizando o algoritmo SAM (*Spectral Angle Mapper*) conseguiram separar as classes:

Campo sujo, Cerrado ralo, Cerrado típico, Cerrado denso, Floresta estacional semidecidual, Floresta estacional decidual, vegetação secundária, cultura agrícola e pastagem cultivada.

Um estudo sobre caracterização de formações savânicas, florestais e áreas de cerrado convertida pela ação antrópica, utilizando séries temporais de NDVI e EVI do sensor MODIS, indicaram que os valores de NDVI foram maiores que os valores de EVI e que existe relação com as estações do ano. As formações florestais foram as que apresentaram maiores valores de NDVI e EVI, possuindo ainda as menores variações entre as estações. As áreas convertidas apresentaram os menores valores dos dois índices, tendo grande queda em seus valores no início da estação seca, provavelmente por se tratar do período de colheita. O estudo enfatiza que as séries temporais NDVI e EVI são úteis na diferenciação dos tipos de vegetação (COUTO JUNIOR et al., 2011).

Lohmann et al., (2009) analisaram e compararam a evolução temporal da reflectância da superfície terrestre, a partir de NDVI da cultura da soja na safra 2007/2008 no estado do Paraná e verificaram que a utilização de imagens de NDVI possibilitou a geração de máscaras de plantios, bem como o monitoramento do desenvolvimento fenológico ao longo das safras e foi possível obter avaliações qualitativas e quantitativas dos plantios, anteriormente às colheitas.

Aguiar et al., (2008) desenvolveram um método operacional usando imagens do sensor MODIS para periodicamente estimar a área colhida de cana-de-açúcar durante seu período de oito meses de colheita (abril a novembro) no Estado de São Paulo. O resultado do estudo mostrou que foi possível monitorar a colheita de cana-de-açúcar com imagens do sensor MODIS. A aplicação do limiar de perda de biomassa sobre as composições MODIS resultou nas melhores estimativas da área de cana colhida. As composições MODIS se mostram apropriadas para detecção do corte da cana, pois evidenciaram as áreas colhidas ao longo da safra, além de aumentarem a precisão das estimativas das datas de colheita.

Para detectar padrões de cultivo para a região do Cerrado, em Comodoro no Mato Grosso, utilizando séries temporais, Galford et al., (2008) usaram imagens EVI do MODIS no período de 5 anos (2000 - 2005). Durante o período de estudo, verificaram um aumento de áreas de cultivo de 3.281 km². A intensificação das plantações se tornou evidentes com o aumento das culturas principalmente nos anos de 2003 e 2004.

Em outro estudo, cujo objetivo foi obter o monitoramento e a estimativa da produtividade de café, a partir de um modelo agrometeorológico-espectral e verificar o comportamento dos índices de vegetação NDVI e EVI do sensor MODIS para diferentes lavouras de café, observou-se que havia uma diferença no comportamento dos índices de

vegetação entre as lavouras de café em produção, formação e podas. Os dois índices apresentaram coerência aos estádios fenológicos da cultura, e o modelo mostrou-se eficiente na estimativa da produtividade em escala regional. Com base no modelo agrometeorológico-espectral foi possível estimar a produtividade e determinar a quebra de produção acumulada por período fenológico da cultura (ROSA et al., 2010).

Para mapear automaticamente os principais tipos de culturas agrícolas em três regiões que cobrem partes do Kansas e Nebraska nos EUA e um terceiro, no noroeste da Turquia, Ozdogan (2010) desenvolveu um algoritmo de processamento de sinal supervisionado chamado Análise de Componentes Independentes (ICA) para decompor os dados temporalmente do MODIS. O estudo demonstrou que o método ICA aplicado a dados MODIS de séries temporais é uma ferramenta viável para mapeamento automatizado de tipos de culturas em escala regional, pois, os padrões gerais, em nível de paisagem de cultivo de inverno e verão retratados nos mapas da ICA são semelhantes àsquelas feitas a partir de observações realizadas em imagens Landsat em ambos os locais.

Couto Júnior et al., (2012) analisaram uma série temporal do sensor MODIS para descrição das sucessões soja-milho em milho-soja e sistema de plantio direto. Os dados temporais apresentam respostas sazonais para o cultivo de sucessão de milho-soja e soja-milho, mostrando-se consistentes com a fenologia das culturas. Os perfis temporais tratados no estudo demonstraram os comportamentos ao longo do tempo, sendo que a soja como primeira cultura apresentou maiores valores para todos os índices avaliados.

Em duas regiões agrícolas representativas na China, uma em torno de Tongzhou, Pequim e outros localizados ao redor Shuyang, Jiangsu, que cultivam trigo de inverno, foi desenvolvido um índice de proporção da fenologia da cultura (CPPI) para: estimar a fração de cultivos individuais usando os dados da série temporal do MODIS-EVI, examinar a influência de diferentes resoluções espaciais sobre o desempenho do índice validado em todo local de estudo e avaliar o potencial de expansão do CPPI para estimativa de área cultivada em larga escala. Este estudo apresentou bom desempenho utilizando as informações fenológicas de crescimento e senescência do trigo de inverno em regiões tipicamente dominantes, um com menor intensidade de cultivo, com tamanho de campos menores e outro com maior intensidade de cultivo, com campos maiores (PAN et al., 2012).

Em uma pesquisa realizada em três locais no Centro de Investigação de Desenvolvimento Agrário, Nebraska-EUA, em oito estações de crescimento de 2001 a 2008, Peng et al., (2013) desenvolveram um modelo para estimar a produtividade primária bruta (GPP) em áreas irrigadas com cultivo de milho e soja de sequeiro, utilizando dados do

MODIS de 250 m. O milho e a soja são tipos de culturas diferentes em estruturas foliares e arquitetura de dossel, e ainda apresentam diferentes tipos de manejo e condições climáticas. Porém, o estudo mostrou que o modelo desenvolvido foi capaz de detectar com precisão a variação do GPP e em culturas onde o conteúdo de clorofila total estava intimamente ligado a dinâmica sazonal de GPP.

Estes estudos corroboram o sucesso das tentativas iniciais para mapear automaticamente as distribuições de culturas em grandes áreas ao redor do mundo usando dados MODIS. E são vistos como um passo inicial para o desenvolvimento de ferramentas de monitoramento de colheitas baseados em MODIS. Comprovam também que o mapeamento e a geração de informações a respeito do desenvolvimento das lavouras de grãos são fundamentais para aprimorar as estimativas de áreas e rendimento das mesmas.

1.4. FENOLOGIA DA VEGETAÇÃO

Segundo Kuplich et al., (2013) a identificação da fenologia em ciclos ou padrões sazonais em séries temporais de dados de vegetação possibilita observar anomalias e efeitos de mudanças climáticas ou ambientais nesses ciclos. Obter informações sobre a dinâmica fenológica dos ecossistemas e sistemas agrícolas pode anteceder ações de conservação e manejo, possibilitar e prever ritmos de crescimento e produção, e aumentar as chances de sucesso dessas ações.

Vários estudos sobre a fenologia da vegetação têm sido realizados, nos últimos anos, com a utilização de imagens de satélite. Isto tem possibilitado monitorar os processos da dinâmica da vegetação, suas mudanças e os impactos. Em geral, os estudos utilizam índices de vegetação como o NDVI e o EVI. Moulin et al., (1997) avaliaram os principais estágios fenológicos da vegetação global por meio de utilização do NDVI proveniente de imagens AVHRR e evidenciaram que as medições por satélite proporcionam monitorar em escala global a fenologia.

Nesta mesma linha de pesquisa, Zhang et al., (2001) desenvolveram um algoritmo de detecção de curvatura máxima para medir as datas de transição fenológica de cada ciclo da vegetação. Este método foi aplicado, com base em dados MODIS e AVHRR, para constatar a fenologia da vegetação de extensões regionais para globais. Este procedimento forneceu dados de séries temporais mais confiáveis.

1.5. ESTIMATIVA DE ESTÁGIOS FENOLÓGICOS DE CULTURAS AGRÍCOLAS

Um trabalho precursor no uso de geotecnologias para monitorar os ciclos fenológicos de culturas agrícolas foi concretizado por Tucker et al., (1979). Por meio da radiometria de campo, analisaram o crescimento e o desenvolvimento do milho e da soja, a partir da relação entre a resposta espectral do vermelho e do infravermelho próximo, do índice de vegetação de razão simples e das fenologias das culturas. Os autores conseguiram identificar 5 estágios fenológicos das culturas e apontaram que a ferramenta de sensoriamento remoto poderia auxiliar na previsão de safras das culturas.

Estudos mostraram que o índice espectral para todo o ciclo de crescimento de uma cultura agrícola sugeria uma curva contínua que mostrasse um aumento distinto e uma queda, com uma curva em forma de sino. Badhwar (1980) recomendou uma função analítica para descrever esse comportamento temporal do índice espectral, enquanto que Henderson & Badhwar (1984), usando dados remotos adquiridos do sistema de satélite Landsat, iniciaram estudos de modelos de simulação de crescimento de culturas e modelos da fenologia em nível de campo em uma grande área geográfica.

Outro estudo utilizou a composição semanal de vários índices de vegetação de dados do visível e do infravermelho próximo do NOAA / AVHRR (National Oceanic and Atmospheric Administration / Advanced Very High Resolution Radiometer) e identificou a ocorrência do estágio de espigamento do milho e indicou o uso deste método quando as informações terrestres confiáveis sobre o desenvolvimento fenológico da cultura não estiverem disponíveis (GALLO & FLESCHE, 1989). Posteriormente, foi desenvolvida uma equação semiempírica usada para representar as variações temporais no NDVI, calculados a partir de medições do NOAA / AVHRR. O modelo estimou os estágios fenológicos das culturas de verão e inverno. Foi destacado que os pontos de inflexão das curvas ascendente e descendente transmitiam informações importantes sobre o desenvolvimento da cultura e ressaltaram que o modelo poderia estimar o rendimento das culturas em escala regional (FISCHER, 1994).

Segundo Sakamoto et al., (2005) as informações da fenologia são essenciais para avaliar a produtividade das culturas e o seu manejo. Por isso, desenvolveram um método para determinar remotamente os estágios fenológicos do arroz em casca. Utilizaram as imagens EVI do MODIS, onde aplicaram técnicas de análise harmônica e de Wavelet para filtrar os dados antes de estimarem a fenologia do arroz. Nesse trabalho, além de indicarem o melhor filtro para tratar os dados, mostraram que foi possível estimar os estágios fenológicos das culturas de arroz. Mais tarde, Sakamoto et al., (2010), utilizaram o WDRVI (Wide Dynamic

Range Vegetation Index) derivados do MODIS para detectar os estágios fenológicos da soja e do milho a partir da série temporal. Nesse estudo, estimaram com precisão a data de quatro estágios fenológicos do milho: início do estágio vegetativo, espigamento, enchimento de grãos e maturidade, e da soja: início do estágio vegetativo, enchimento de grão, grão completo e início da maturação.

Esses estudos confirmam a importância de utilizar o sensoriamento remoto para estimar a produtividade agrícola e detectar as mudanças dos ciclos fenológicos. Dessa forma, se torna cada vez mais relevante o desenvolvimento de tecnologias capazes de refinar as observações ao longo de cada cultura agrícola.

1.6. ANÁLISE DE SÉRIES TEMPORAIS

As séries temporais de imagens de satélite são importantes, pois permitem a produção de coleção de dados que ajudam a desenvolver estudos que necessitam de um enfoque temporal, como por exemplo: monitoramento de safras agrícolas e ambiental, caracterização da fenologia da vegetação, mudança sazonal e detecção de mudanças no uso e cobertura do solo.

Diversas pesquisas utilizaram dados temporais com o escopo de estimar de forma mais precisa as alterações que ocorrem na paisagem em curto intervalo de tempo, por meio de eventos naturais ou antrópicos. Galford et al., (2008) com o objetivo de detectar padrões de cultivo no Cerrado em Mato Grosso, utilizaram uma série temporal de imagens EVI do MODIS no período de 5 anos, e identificaram a intensificação e o aumento dos cultivos agrícolas na região.

Para o monitoramento agrícola e previsão de safras, pesquisas têm sido desenvolvidas com a análise de séries temporais, com o objetivo de proporcionar uma maior agilidade na aquisição de dados e menor custo em relação a métodos tradicionais de censos e pesquisas. Desta forma, Araújo et al., (2011) mapearam culturas de verão no Estado do Paraná por meio de séries temporais, identificando áreas importantes de ocupação da cultura de verão no Estado, além de notar possíveis diferenças na época de plantio da safra em algumas regiões do Estado. Gusso et al., (2012) desenvolveram uma metodologia para estimar a área plantada de soja no Estado do Rio Grande do Sul com base na análise de séries temporais e as estimativas de área plantada concordaram com as estatísticas agrícolas oficiais. Arvor et al., (2011) utilizando séries temporais, apresentaram uma metodologia para mapeamento das principais culturas agrícolas em escala regional no Estado do Mato Grosso e a metodologia desenvolvida

por eles confirmou a eficiência do MODIS na obtenção de estatísticas agrícolas em níveis estadual e municipal.

1.7. SUAVIZAÇÃO DAS SÉRIES TEMPORAIS

Objetivando preencher as lacunas com informações frequentes dos processos e eventos no tempo e no espaço, o monitoramento da dinâmica do uso e cobertura do solo tem impulsionado o avanço na utilização dos dados de satélites da série temporal. Entretanto, estes dados de satélites podem estar sujeitos a interferências de diversos fatores relacionados com presença de ruídos. Os ruídos estão associados à presença de nuvens, sombras e variações da atmosfera. Diante disso, trabalhar com estes dados requer um processamento de suavização a partir de técnicas de filtragens para a eliminação destes ruídos.

Vários métodos de suavização foram desenvolvidos para a extração desses ruídos, dentre eles destacam-se o Savitzky–Golay (SAVITZKY & GOLAY, 1964), o Fourier (AZZALI & MENENTI, 2000), mediana (Ataman et al., 1981), função duplo-logística (BECK et al., 2006), gaussiano assimétrico (JÖNSSON & EKLUNDH, 2002), Wavelets (GRAPS, 1995), entre outros.

O filtro Savitzky-Golay (SAVITZKY & GOLAY, 1964) é um dos métodos de suavização de sinais analíticos mais utilizados, minimiza os ruídos através da interpolação linear, baseada em média móvel ponderada, cujo valor da janela é pré-estabelecido conforme a natureza do ruído. A principal vantagem deste filtro é que ele tende a preservar as características iniciais tais como valores mínimos e máximos, bem como a largura da curva temporal. Chen et al., (2004) indicam este método por ele ser mais eficaz na obtenção de uma elevada qualidade das séries temporais do NDVI.

A transformada de Fourier converte um sinal medido no domínio do tempo de um sinal no domínio da frequência (SAKAMOTO et al., 2005). Neste método cada onda é definida por uma amplitude única e um ângulo de fase, onde o valor de amplitude é metade da altura de uma onda, e o ângulo de fase define o afastamento entre a origem e o pico da onda ao longo do intervalo (JAKUBAUSKAS et al., 2002). A transformada de Fourier foi aplicada sobre séries temporais de produtos NDVI objetivando avaliar as amplitudes das alterações periódicas das respostas da vegetação no tempo (AZZALI & MENENTI, 2000; WAGENSEIL & SAMIMI, 2006; MINGWEI et al., 2008; VICTORIA et al., 2012).

A filtragem de mediana é uma técnica de alisamento não linear, que é conhecido para preservar mudanças bruscas de sinais e por ser particularmente eficaz na remoção de ruído

impulsivo (ATAMAN et al., 1981). Para o tratamento do ruído presente no espectro temporal, um filtro convolucional de mediana que trabalha ao longo do espectro foi desenvolvido por Carvalho Júnior et al., (2008). A partir de uma janela definida pelo usuário, o valor mediano dos dados de entrada é selecionado para substituir o valor do pixel central desta janela (CARVALHO JÚNIOR et al., 2012).

O filtro de duplo-logística foi proposto por Beck et al., (2006) para monitorar o NDVI em latitudes elevadas. Julien & Sobrino, (2009) utilizaram este filtro nos dados NDVI para descrever o inventário do mapeamento global e estudos de monitoramento da evolução anual para o período de 1981 a 2003, com o objetivo de estimar os parâmetros fenológicos da superfície terrestre.

Jönsson & Eklundh (2002; 2004) apresentaram o filtro gaussiano assimétrico para extrair da série temporal de NDVI os parâmetros fenológicos da vegetação. Neste filtro preliminarmente é preciso calcular os parâmetros para ajuste da função local. Por ser um método mais limitado é menos sensível ao ruído e possibilita a utilização dos ajustes mais incisivos nos dados. As funções locais são adaptadas aos dados que se encontram nos intervalos em volta do valor mínimo e máximo.

A transformada de Wavelets são funções matemáticas que reduz os dados em diferentes componentes de frequência, e em seguida, estudam cada componente com uma resolução associada à sua escala. Eles têm vantagens sobre os métodos tradicionais de Fourier em análise de situações físicas em que o sinal contém descontinuidades e pontas agudas. Wavelets foram desenvolvidos de forma independente nos campos da matemática, a física quântica, engenharia elétrica, geologia e sísmica. Intercâmbios entre esses campos durante os últimos dez anos levaram a muitas novas aplicações Wavelets tais como compressão de imagem, turbulência, a visão humana, radar, e previsão de terremotos (GRAPS, 1995). Segundo Rosemback et al., (2009) a transformada de Wavelets é uma ferramenta que possui um grande potencial para a observação da dinâmica sazonal da vegetação e, por conseguinte para observar perturbações no seu comportamento padrão frente a fenômenos climáticos.

1.8 – MÉTODOS DE CLASSIFICAÇÃO

As imagens de satélites possuem uma grande diversidade de alvos complexos. Para extrair essas informações, é necessário o uso de classificações automáticas. Dependendo do algoritmo utilizado neste procedimento, a classificação é denominada supervisionada ou não-supervisionada. Na abordagem supervisionada, o classificador orienta a sua busca de alvos a partir de amostras de treino selecionadas anteriormente com os alvos de interesse da área de

estudo. Na não-supervisionada, não há qualquer conhecimento prévio do classificador sobre os atributos dos alvos pertinentes da área de estudo. Dentro dos algoritmos supervisionados, os mais empregados são: Mínima Distância (MD), Spectral Angle Mapper (SAM), Artificial Neural Network (ANN), Support Vector Machines (SVMs) e vizinho mais próximo (K-NN).

O algoritmo de Mínima Distância utiliza a distância para associar um pixel a uma determinada classe. Na classificação cada pixel será incorporado a um agrupamento, através da análise de similaridade da distância (RICHARDS, 1999). Carvalho Júnior et al., (2009), utilizaram o algoritmo de mínima distância para identificar as fisionomias de Cerrado a partir da classificação de séries temporais NDVI-MODIS. Wang & Tenhunen (2004), aplicaram o Mínima Distância em uma série temporal NDVI, para analisar as precisões de mapeamento da vegetação e fornecer uma classificação atualizada, no Nordeste da China.

O SAM é uma classificação espectral que utiliza um ângulo de n-D para corresponder pixels para fazer a referência espectral. O algoritmo determina a semelhança espectral entre os dois espectros, calculando o ângulo entre os espectros e os tratam como vetores num espaço com dimensionalidade igual ao número de bandas (KRUSE et al., 1993). Arvor et al., (2011), aplicaram três métodos de classificação, entre eles o SAM, para mapear as principais culturas e práticas agrícolas em escala regional para o Estado do Mato Grosso, utilizando imagens EVI MODIS. Borges & Sano (2014), utilizaram o algoritmo SAM nas séries temporais de EVI MODIS para mapear classes de cobertura vegetal natural e antrópica no Oeste da Bahia.

As Redes Neurais artificiais são técnicas de processamento de informações produzidas pela interligação de unidades simples de processamento, denominadas neurônios artificiais, que recebe essa intitulação porque foram criados a partir de um modelo matemático simplificado de um neurônio natural. Elas são dispositivos não-lineares, calculados por determinadas funções matemáticas (HAYKIN, 2001). Panda et al., (2010) investigaram quatro índices espectrais para prever o rendimento das culturas agrícolas usando técnicas de rede neural, em Dakota do Norte, nos Estados Unidos. Atzberger & Rembold (2013) avaliaram o desempenho das redes neurais para mapear a distribuição das culturas de inverno usando séries temporais de imagens AVHRR NDVI, em Toscana, na Itália.

O SVM (Support Vector Machines) é um sistema de classificação derivado da teoria de aprendizagem estatística. Separa as classes com uma superfície de decisão que maximiza a margem entre as classes (CHANG & LIN, 2013). Chen et al., (2012), estudaram as práticas e as áreas de cultivo de arroz no Delta do Mekong, no Vietnã, utilizando os dados NDVI MODIS, aplicando a classificação SVM. Abade et al., (2015), mapearam a vegetação nativa e

exótica na zona de transição de Cerrado-Caatinga no Brasil Central, utilizando os dados da série temporal MODIS-NDVI, e avaliaram quatro classificadores entre eles o SVM.

O método vizinho mais próximo (K-NN) é um classificador não supervisionado e não paramétrico que classifica um novo objeto (classe desconhecida), com a mesma classe do objeto mais próximo no conjunto de amostras de treinamento (COVER & HART, 1967). Franco-Lopez et al., (2001) utilizaram a técnica do vizinho mais próximo (K-NN) para monitoramento florestal, onde destacou a transparência do método, por ser alicerçado unicamente na busca de unidades similares. Da mesma forma, Robertson e King (2011) produziram mapas temáticos do uso e cobertura do solo mais uniformes e precisos no qual compararam uma classificação orientada em pixel e objeto, a partir da técnica do vizinho mais próximo (K-NN).

1.9. ANÁLISE DE ACURÁCIA

O conceito de acurácia parece simples, mas é uma propriedade difícil de medir e expressar. Na cartografia temática a partir de dados de sensoriamento remoto, o termo acurácia geralmente é usado para expressar o grau de correção de um mapa ou a classificação. Um determinado mapa temático com uma classificação pode ser considerado preciso se ele fornece uma representação imparcial da cobertura da terra da região que retrata (FOODY, 2002).

A acurácia de mapeamentos de dados globais é fundamental tanto para estabelecer a exatidão dos produtos para a comunidade científica, quanto para fornecer um retorno dos algoritmos de processamento de dados e modelos orientados, para que os produtos possam ser melhorados (COHEN & JUSTICE, 1999). Diante disso, várias técnicas de análise multivariadas têm sido utilizadas para avaliar a acurácia de classificações do uso e cobertura da terra a partir de imagens digitais, o que ajuda a identificar as variáveis ou fatores que afetam a precisão da classificação (CONGALTON et al., 1983). Congalton (1988a) investigou a correção espacial dos dados de sensoriamento para determinar seus efeitos sobre a avaliação da acurácia dos mapeamentos.

Segundo Foody (2002), muitos métodos de avaliação da exatidão têm sido discutidos na literatura. A mais utilizada, no entanto, pode ser derivada a partir de uma matriz de confusão ou erro. A matriz de confusão é atualmente a parte central dos artigos para tratar da avaliação da acurácia. Como uma tabulação cruzada simples, a matriz de confusão constitui a base sobre que descreve a precisão da classificação e caracteriza os erros, podendo ajudar a refinar a classificação ou estimativa derivada dela. Os dados da amostra de referência são

resumidos numa matriz de erro e várias estatísticas de precisão são calculadas a partir dela. Assim, muitos trabalhos empregaram a matriz de confusão para avaliar a acurácia dos seus dados (CONGALTON et al., 1983; CONGALTON, 1988b, 1991; CONGALTON & GREEN, 1993; STEHMAN, 1997; FOODY, 2002; VAN OORT, 2007).

Com o objetivo de verificar a acurácia das classificações deste trabalho, foi realizada uma avaliação de exatidão por meio do índice Kappa, utilizando como verdade terrestre pontos coletados no campo e mapas de referências. Segundo Antunes & Lingnau (2004) o coeficiente de acurácia Kappa para classes individuais, parece ser o mais indicado na avaliação da classificação de imagens de sensoriamento remoto, por levar em conta os erros comissão e omissão na matriz de erros no seu cálculo. Vários autores já utilizaram este índice para avaliar a acurácia oriunda de dados de sensoriamento remoto (STEHMAN, 1996; GOLFORD et al., 2008; FORKEL et al., 2013; ULLMANN et al., 2014; ABADE et al., 2015; GAO et al., 2015; XU et al., 2015).

Além disso, foi aplicado o teste McNemar (MCNEMAR, 1947) para comparar os resultados das acurácias. O McNemar é um teste não paramétrico e que avalia a significância estatística das diferenças entre as classificações, baseando-se em matrizes de confusão de duas dimensões. Neste teste, a atenção é focalizada na diferença binária entre a destinação das classes corretamente e incorretamente. Este tem sido bastante utilizado na detecção remota (FOODY, 2004; LEEUW et al., 2006; HARTFIELD et al., 2011; ALCANTARA et al., 2012).

Alguns autores confrontaram seus resultados com as estimativas oficiais do IBGE e verificaram concordância entre os dados observados (RIZZI & RUDORFF, 2005; LAMPARELLI et al., 2008; ARAÚJO et al., 2011; GUSSO et al., 2012; VICTORIA et al., 2012). Sendo assim, as áreas agrícolas mapeadas com o uso do MODIS poderão ser calculadas e comparadas com os dados obtidos do IBGE.

Para avaliar a precisão da classificação é importante salientar os elementos que incluem a definição do tamanho da amostra, bem como a especificação e utilização de uma medida de precisão adequada para a aplicação. Segundo Foody (2002) o tamanho das amostras deve ser escolhido com cuidado e ser suficiente para proporcionar uma base representativa e útil para avaliação da acurácia e deve ser grande o suficiente para garantir que todas as classes sejam adequadamente representadas.

1.10. SISTEMAS DE MONITORAMENTO GLOBAL DA AGRICULTURA

Pittman et al., (2010) analisaram a capacidade do NDVI do sensor MODIS para mapear áreas cultiváveis em extensão global e avaliaram os resultados para os principais países e regiões de produção de culturas agrícolas. Os resultados revelaram que o MODIS é uma fonte de dados apropriada para o mapeamento das terras agrícolas em escala global. Whitcraft et al., (2015), utilizaram uma série temporal de imagens NDVI do MODIS de 8 dias, para identificar a transição das datas fenológicas do início, período de pico e final de safra da agricultura global, no período de 2001 a 2010. Este estudo forneceu uma estratégia para a aquisição de imagens para o monitoramento agrícola mundial.

Já existe uma série de sistemas de monitoramento agrícola internacionais, que desempenham um papel fundamental nos processos de tomada de decisão e que regem a ajuda alimentar e de produtos agrícolas no mercado global. Alguns destes sistemas utilizam uma combinação de observações oriundas de imagens de satélite (BECKER-RESHEF et al., 2009; PITTMAN et al., 2010).

Entre estes sistemas de monitoramento podemos destacar os globais como o Global Information and Early Warning System (GIEWS) da FAO (Food and Agriculture Organization), o Monitoring Agriculture by Remote Sensing (MARS) da União Européia, o Global Agriculture Monitoring (GLAM) dos Estados Unidos, o Crop Watch System (CAS) da China e o Group on Earth Observations Global Agricultural Monitoring (GEOGLAM) do G20. E ainda os sistemas de monitoramento agrícolas nacionais como: Geosafras do Brasil e o National Agricultural Statistics Service (NASS) dos Estados Unidos (JUSTICE & BECKER-RESHEF, 2007; BECKER-RESHEF et al., 2010).

Segundo Becker-Reshef et al., (2010) a necessidade de monitorar a agricultura global continua a crescer em face de antecipados eventos climáticos mais frequentes e extremos, como inundações e secas, as crescentes necessidades de energia e o aumento da demanda por alimentos impulsionado pelo crescimento populacional e desenvolvimento econômico. Dessa forma, se torna cada vez mais importante a busca por métodos e tecnologias voltadas para o monitoramento da produção agrícola em escala regional, nacional e global.

1.11. AGRICULTURA NAS ÁREAS DE ESTUDO

A agricultura brasileira cresceu bastante nas últimas décadas, estabelecendo o Brasil, como um dos principais produtores agrícolas do mundo. Historicamente, a produção agrícola no país iniciou-se na região Nordeste, com a cana-de-açúcar. Com a introdução de novas tecnologias no processo produtivo, as cinco regiões brasileiras, tiveram um aumento

expressivo na produção agrícola, porém foi mais intenso nas regiões Sul, Sudeste e Centro-Oeste (CASTRO, 2014).

As três áreas agrícolas analisadas neste trabalho (**Figura 1.2**), estão localizadas na Região Centro-Oeste e Sul do país, e foram selecionadas devido a dados de campo pré-existent e cedidos pelos órgãos responsáveis por estes. A Embrapa Agrossilvipastoril e a Associação dos Produtores de Soja e Milho do Estado do Mato Grosso (APROSOJA), além de fornecerem dados de campo sobre as culturas agrícolas de soja, milho e algodão, também deram apoio para um trabalho de campo realizado na região. Da mesma forma, a Companhia Nacional de Abastecimento (Conab) cedeu dados de campo sobre a cultura do trigo no Rio Grande do Sul e a cultura do arroz em Santa Catarina, e também apoiou um trabalho de campo realizado no Estado de Mato Grosso.

Atualmente, as regiões Centro-Oeste e Sul, são responsáveis por cerca de 78,40% da produção nacional de grãos (IBGE, 2015).

A evolução histórica da região Sul, é explicada pela composição estrutural da economia. Até as três primeiras décadas do século XX, predominou a produção agrícola familiar, derivada das atividades em propriedades de imigrantes de origem europeia (MONTIBELLER FILHO & GARGIONI, 2014).

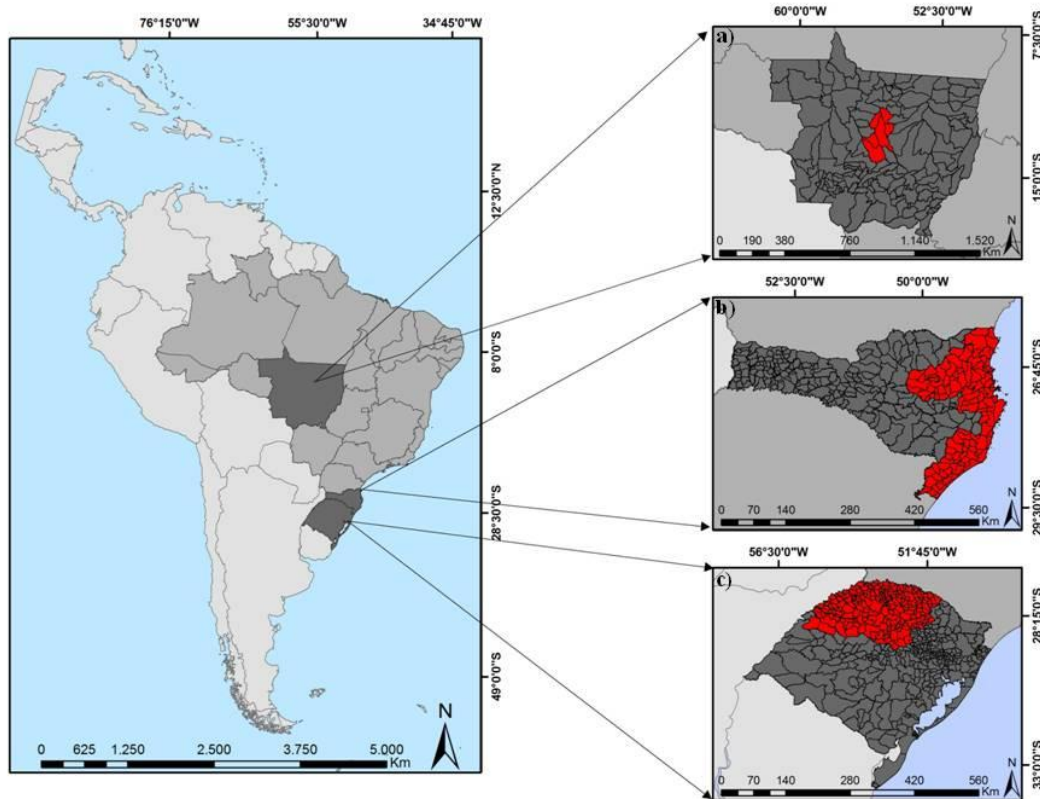


Figura 1. 2. Localização da três áreas de estudo

A agricultura desenvolvida na região Sul do Brasil, possui diversas características próprias que a diferencia das outras regiões brasileiras, como a estrutura agrária menos concentrada, os agricultores familiares estão dispersos por todo o território e em geral utilizam mais insumos modernos de produção e obtêm melhor rentabilidade financeira. Entre os principais produtos da região destaca-se a produção de soja e milho (CASTRO, 2014).

Na região Centro-Oeste, os fatores que contribuíram para o desenvolvimento da agricultura foram a política de incentivos fiscais, o grande percentual de áreas com relevo plano facilitando a mecanização da agricultura, a localização central e malha viária, que promove o escoamento da produção, além de condições climáticas favoráveis.

Helfand & Rezende (2000) destacam que a produção de grãos no Brasil atingiu níveis sem precedentes nos anos 1990. Na região Centro-Oeste cresceu rapidamente, em torno de aproximadamente 585% nos períodos de 1973 a 1999, na região Sul e Nordeste no mesmo período cresceram em torno de 122,15% e 63,15%, respectivamente, enquanto no Sudeste cresceu em torno de 45,12% (Tabela 1.1).

Tabela 1. 1- Área e Produção de Grãos por Região — 1973/99

(Em Milhões de ha e t)

Anos	Norte		Nordeste		Sudeste		Sul		Centro-Oeste	
	Área	Produção	Área	Produção	Área	Produção	Área	Produção	Área	Produção
1973	0,2	0,2	5,6	3,8	6,3	8,2	11,3	16,7	2,6	3,5
1974	0,3	0,3	5,5	3,3	6,3	8,7	13,9	21,4	2,9	3,8
1975	0,3	0,4	5,9	3,8	6,4	8,0	14,8	23,1	3,3	4,2
1976	0,3	0,4	5,2	2,8	6,6	9,1	16,5	27,2	4,2	5,2
1977	0,4	0,4	6,2	4,1	6,5	8,8	16,6	27,4	4,4	5,8
1978	0,4	0,4	6,1	3,9	6,4	7,5	16,1	20,4	4,4	4,2
1979	0,4	0,5	5,9	3,5	5,9	8,6	17,6	21,9	4,5	6,1
1980	0,5	0,7	5,8	3,0	6,4	9,9	17,6	30,5	5,2	7,5
1981	0,6	0,7	5,8	2,0	6,7	10,1	16,1	30,9	5,2	6,7
1982	0,6	0,8	7,5	4,0	6,9	11,0	16,6	26,8	5,5	8,1
1983	0,4	0,5	4,1	1,3	6,2	10,5	15,3	26,0	5,3	8,2
1984	0,7	0,9	6,7	4,4	6,5	9,7	15,7	28,5	5,9	8,8
1985	0,7	0,8	6,9	3,9	6,6	11,5	16,5	31,8	6,2	10,8
1986	0,8	1,0	8,2	5,5	6,8	11,9	16,0	24,3	6,9	11,6
1987	0,8	0,9	6,7	2,4	6,8	12,4	16,8	34,3	7,6	13,9
1988	1,0	1,2	8,5	6,0	6,7	12,8	16,7	30,8	7,8	14,7
1989	1,0	1,3	8,1	5,1	6,6	13,0	16,6	34,6	8,2	17,4
1990	0,8	0,9	5,9	2,5	6,2	9,4	15,8	31,5	6,9	11,4
1991	0,9	1,1	7,4	5,4	6,2	12,8	15,1	23,3	6,1	13,3
1992	0,9	1,1	7,0	3,5	6,2	12,5	15,3	35,6	6,7	14,6
1993	1,1	1,4	4,5	3,3	5,8	12,3	15,0	36,5	6,8	15,1
1994	1,2	1,6	8,2	7,1	5,9	11,9	14,7	36,0	7,7	18,3
1995	1,2	1,7	7,9	6,6	5,6	12,5	14,6	39,5	7,8	18,8
1996	1,3	1,8	7,8	6,7	5,1	11,2	14,5	34,7	7,0	18,2
1997	1,3	1,8	7,8	6,6	5,0	12,4	14,4	36,2	7,2	19,3
1998	1,4	2,0	4,4	3,8	4,7	11,3	14,3	39,7	7,9	20,5
1999	1,5	2,3	5,8	6,2	4,9	14,1	14,1	37,1	8,5	24,0

Fonte: IBGE – Produção Agrícola Municipal (período 1973/96) e Levantamento Sistemático da Produção Agrícola (1997/99). (Apud Helfand & Rezende, 2000, p. 3).

Rezende (2002) enfatiza que a alta produtividade na região Centro-Oeste se deu por meio das inovações tecnológicas desenvolvidas para agricultura brasileira que solucionou os problemas de baixa fertilidade e elevada acidez dos solos. Além disso, descobertas de novas variedades de sementes e da melhoria genética, além de outros resultados de pesquisas agrícolas que tornaram mais rentáveis a utilização das terras do Cerrado.

Segundo o IBGE (2016), o Estado do Mato Grosso é o maior produtor nacional de grãos. Sendo seus principais produtos a soja, o milho e o algodão. A soja representa 64% da produção do Estado, seguido pelo milho com 28% e o algodão 4% (**Figura 1.3**). Este Estado é dividido em cinco mesorregiões geográficas, sendo que a mesorregião Norte é a maior produtora do Estado, com 65% da produção. O algodão nessa mesorregião representa 61% da produção do Estado, o milho 69% e a soja 64% (**Figura 1.4**). Na área de estudo referente ao capítulo 2, a qual composta por seis municípios da mesorregião Norte: Ipiranga do Norte, Lucas do Rio Verde, Nova Mutum, Sinop, Sorriso e Vera (**Figura 1.2a**), apresenta 21% da produção total de grãos do Estado. Sendo que a produção do algodão, milho e soja, representam 11%, 28% e 19% da produção do Estado, respectivamente (**Figura 1.5**).

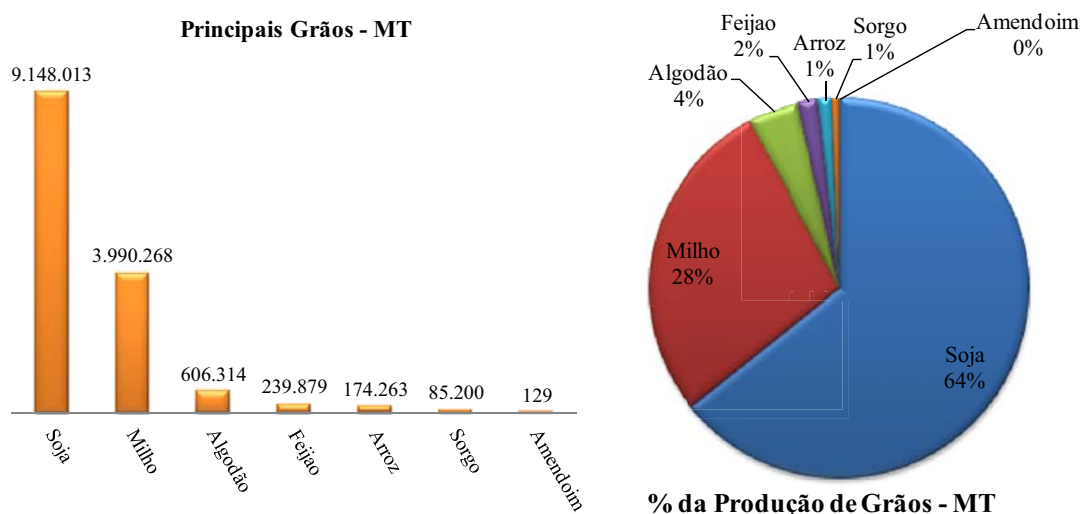


Figura 1.3- Estimativa da produção de grãos no Estado do Mato Grosso

Fonte: IBGE, 2016.

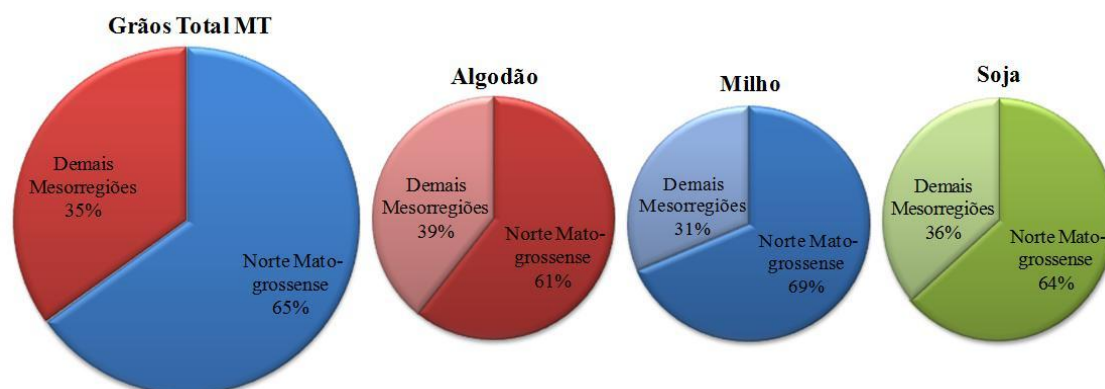


Figura 1. 4- Estimativa da produção dos principais grãos distribuídos por mesorregiões

Fonte: IBGE, 2016.

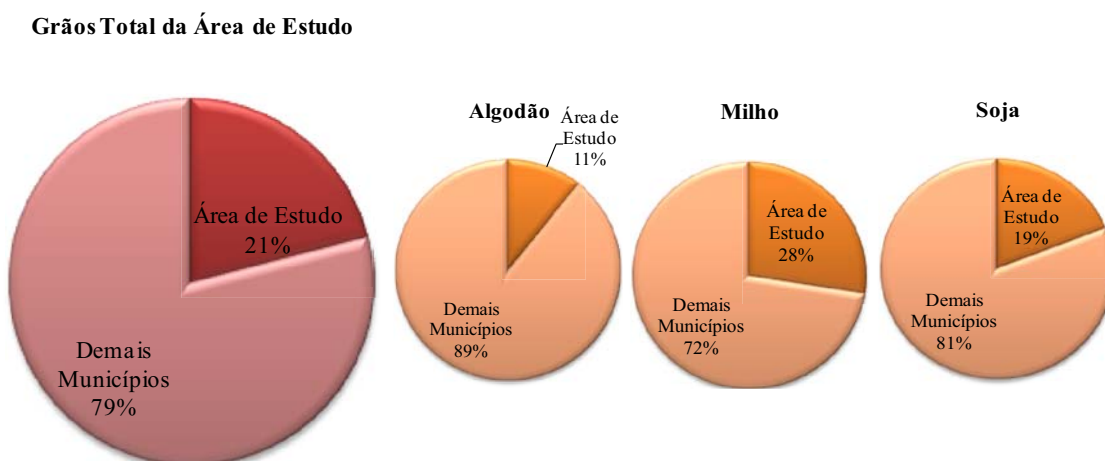


Figura 1. 5- Estimativa da produção dos principais grãos distribuídos na área de estudo

Fonte: IBGE, 2016.

A soja, o arroz, o trigo e o milho são as principais culturas agrícolas do Estado do Rio Grande do Sul. Sendo que a soja representa 65% da produção agrícola do Estado, o arroz cerca de 13%, trigo e milho representam cerca de 9% cada (**Figura 1.6**). O Estado é dividido em sete mesorregiões geográficas, sendo que a mesorregião Noroeste Rio-grandense é a maior produtora do Estado, com 51% da produção. As principais culturas dessa mesorregião são o trigo, a soja e o milho que correspondem 79%, 56% e 47%, respectivamente, da produção total do Estado (**Figura 1.7**). A área de estudo referente ao capítulo 3 é toda a mesorregião Noroeste Rio-grandense (**Figura 1.2c**), que é composta por 216 municípios (IBGE, 2016).

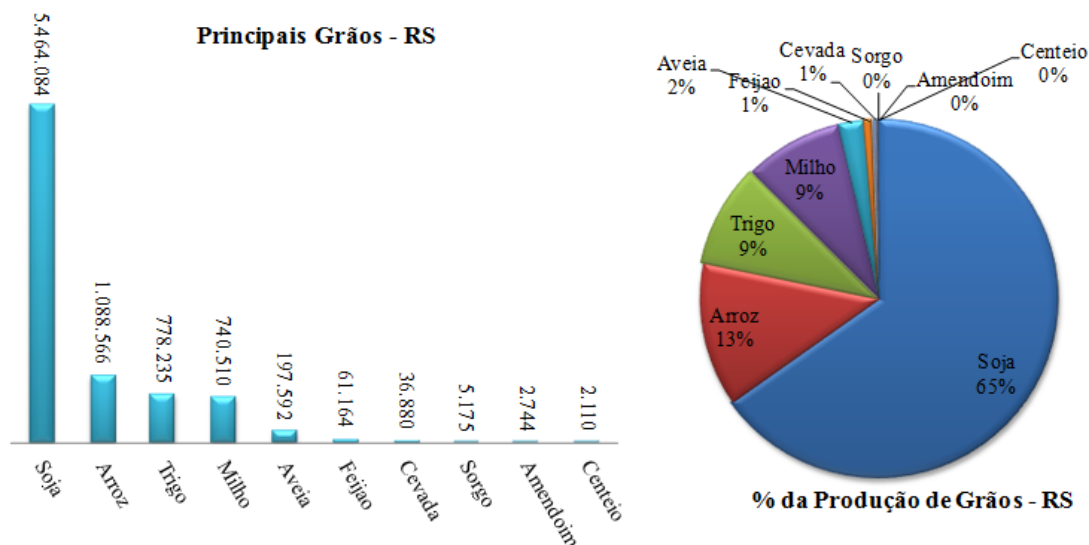


Figura 1. 6- Estimativa da produção de grãos no Estado do Rio Grande do Sul

Fonte: IBGE, 2016.

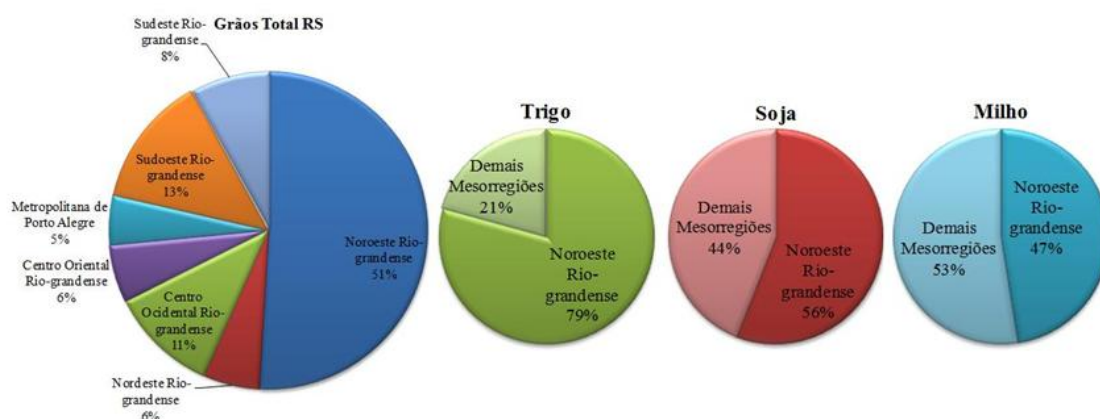


Figura 1. 7- Estimativa da produção dos principais grãos distribuídos por mesorregiões

Fonte: IBGE, 2016.

No Estado de Santa Catarina, as principais culturas agrícolas são a soja, o milho e o arroz. Dentre essas culturas, a soja se destaca, representando 50%. O milho aparece em segundo lugar, e corresponde à 27% da produção agrícola do estado e o arroz 11% (**Figura 1.8**) O Estado é dividido em seis mesorregiões geográficas, da qual a mesorregião Oeste Catarinense é a maior produtora do Estado, com 47% da produção. As principais culturas dessa mesorregião são a soja e o milho que possuem 51%, 57%, respectivamente, da produção total do Estado (**Figura 1.9**). A mesorregião com a maior produção de arroz é a Sul Catarinense, com 63%. A área de estudo referente ao capítulo 4 é composta por nove microrregiões: Araranguá, Blumenau, Criciúma, Florianópolis, Itajaí, Joinville, Rio do Sul, Tijucas e Tubarão (**Figura 1.2b**). Sendo que a Microrregião de Araranguá se destaca com 35% da produção total do Estado (**Figura 1.10**) (IBGE, 2016).

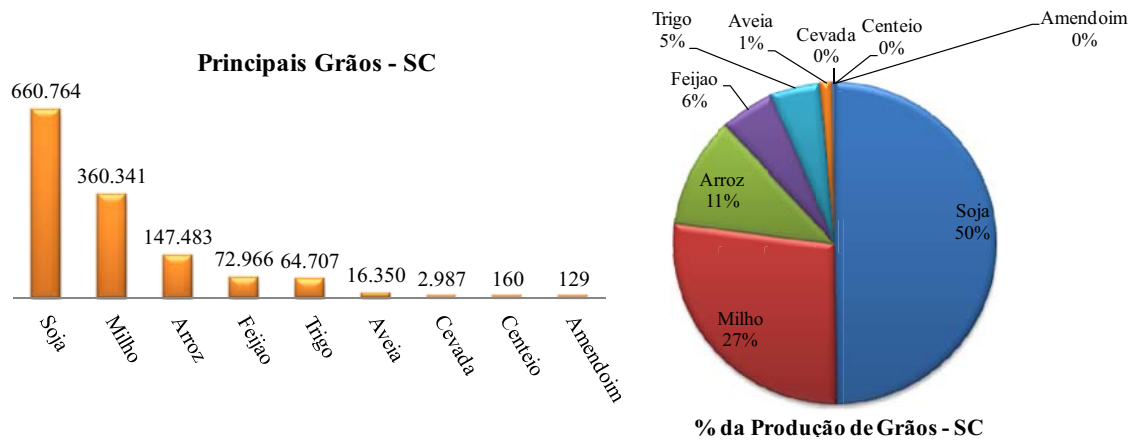


Figura 1. 8- Estimativa da produção de grãos no Estado de Santa Catarina

Fonte: IBGE, 2016.

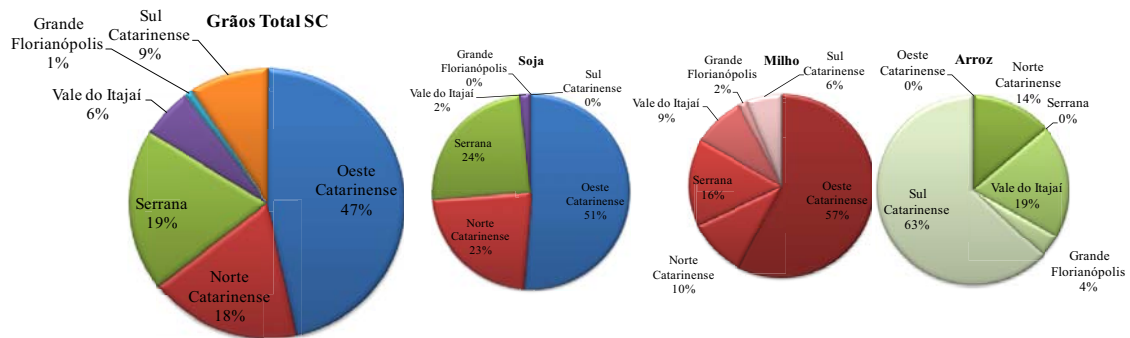


Figura 1. 9- Estimativa da produção dos principais grãos distribuídos por mesorregiões

Fonte: IBGE, 2016.

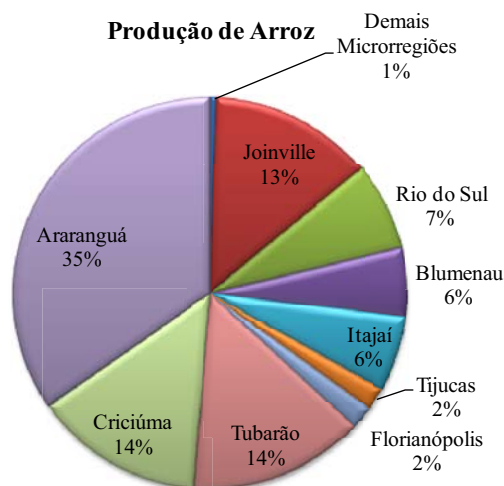


Figura 1. 10- Produção de arroz na área de estudo

Fonte: IBGE, 2016.

1.12. ORGANIZAÇÃO DA TESE

A presente tese está estruturada na forma de artigos. No primeiro capítulo foram apresentados os principais problemas das estimativas de safras no país, destacando a expansão e ocupação dos solos para agricultura nas Regiões Centro-Oeste, Sul e Sudeste. Também foi exposto o uso do sensor MODIS para o monitoramento de áreas naturais e agrícolas, a análise de séries temporais para estudar os diferentes tipos de culturas agrícolas e os principais sistemas de monitoramento global da agricultura.

A parte de desenvolvimento consta com a elaboração de três artigos científicos:

✓ Primeiro Artigo – Faz uma abordagem sobre o uso de séries temporais do sensor MODIS comparando os métodos de classificação SVM e Redes Neurais, em seis municípios produtores de soja, milho e algodão, no Estado do Mato Grosso.

✓ Segundo Artigo – Utiliza-se o método de classificação do vizinho mais próximo (K-NN) a partir de duas métricas: similaridade e distância euclidiana, para distinguir a cultura do trigo na Mesorregião Noroeste do Rio Grande do Sul.

✓ Terceiro Artigo -. Identifica-se a cultura do arroz, usando o método de classificação do vizinho mais próximo (K-NN) a partir de duas métricas: similaridade e distância euclidiana, no Estado de Santa Catarina.

A **Figura 1.11** apresenta a organização da tese.

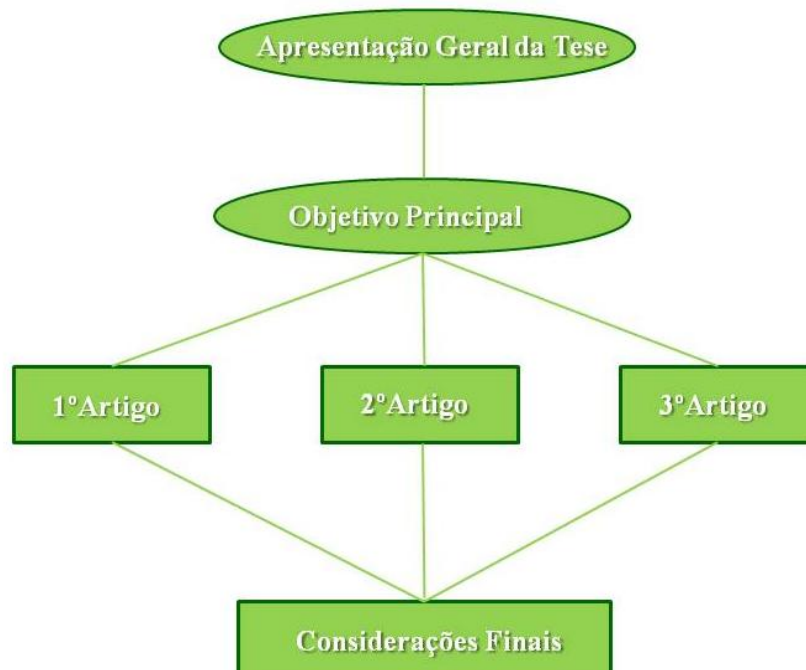


Figura 1. 11- Estrutura da tese

No último capítulo, serão apresentadas as considerações finais.

REFERENCIAS BIBLIOGRÁFICAS

- ABADE, N. A.; DE CARVALHO, O. A.; GUIMARÃES, R. F.; DE OLIVEIRA, S. N. Comparative analysis of MODIS time-series classification using support vector machines and methods based upon distance and similarity measures in the brazilian cerrado-caatinga boundary. **Remote Sensing**, v. 7, n. 9, p. 12160–12191, 2015.
- AGUIAR, D. A. De; FRIEDRICH, B.; RUDORFF, T.; RIZZI, R.; SHIMABUKURO, Y. E. Monitoramento da colheita da cana-de-açúcar por meio de imagens MODIS. **Revista Brasileira de Cartografia**, v. 60, p. 375–383, 2008.
- ALCANTARA, C.; KUEMMERLE, T.; PRISHCHEPOV, A. V.; RADELOFF, V. C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. **Remote Sensing of Environment**, v. 124, p. 334–347, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.rse.2012.05.019>>.
- AMARAL, G. M.; SOUZA, A. L. F.; PIFFER, T. R. O.; SOUSA, A. O.; SANTOS, E. C.; SOUZA, L. M. M.; OLIVEIRA, C. C. Uso da geotecnologia para o mapeamento da cultura do café na microrregião de Ariquemes no Estado de Rondônia. In: XV Simpósio Brasileiro de Sensoriamento Remoto, Curitiba - PR. Curitiba - PR: 2011.
- ANTUNES, A. F. B.; LINGNAU, C. Analise de acurácia de mapa de uso do solo oriundo de classificação de imagem de alta resolução. **Boletim de Ciencias Geodesicas (Impresso)**, v. 4, p. 101–113, 2004.
- ARAÚJO, G. K. D.; ROCHA, J. V.; LAMPARELLI, R. A. C.; ROCHA, A. M. Mapping of summer crops in the State of Paraná, Brazil, through the 10-day spot vegetation NDVI composites. **Engenharia Agrícola**, v. 31, n. 4, p. 760–770, 2011.
- ARVOR, D.; JONATHAN, M.; MEIRELLES, M. S. P.; DUBREUIL, V.; DURIEUX, L. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. **International Journal of Remote Sensing**, v. 32, n. 22, p. 7847–7871, 2011.
- ATAMAN, E.; AATRE, V.; WONG, K. Some statistical properties of median filters. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, v. 29, n. 5, p. 1073–1075, 1981.
- ATZBERGER, C.; REMBOLD, F. Mapping the spatial distribution of winter crops at sub-pixel level using AVHRR NDVI time series and neural nets. **Remote Sensing**, v. 5, n. 3, p. 1335–1354, 2013.
- AZZALI, S.; MENENTI, M. Mapping vegetation-soil-climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. **International Journal of Remote Sensing**, v. 21, n. 5, p. 973–996, 2000.
- BADHWAR, G. D. Crop emergence date determination from spectral data. **Photogrammetric Engineering & Remote Sensing**, v. 46, n. 3, p. 369–377, 1980.
- BARROS, G. S. C.; ADAMI, A. C. O. **Faturamento do agronegócio supera US\$ 100 bilhões em 12 meses.** Disponível em: <http://www.cepea.esalq.usp.br/comunicacao/Cepea_IndicesExport_2tri2013.doc>. Acesso em: 23 jul. 2015.
- BECK, P. S. A.; ATZBERGER, C.; HØGDA, K. A.; JOHANSEN, B.; SKIDMORE, A. K. Improved monitoring of vegetation dynamics at very high latitudes: A new method using

MODIS NDVI. **Remote Sensing of Environment**, v. 100, n. 3, p. 321–334, 2006.

BECKER-RESHEF, I.; JUSTICE, C.; DOORN, B.; REYNOLDS, C.; ANYAMBA, A.; TUCKER, C. J. NASA's contribution to the Group on Earth Observations (GEO) Global Agricultural Monitoring System of Systems. **NASA Earth Observer**, v. 21, p. 24–29, 2009.

BECKER-RESHEF, I.; JUSTICE, C.; SULLIVAN, M.; VERMOTE, E.; TUCKER, C.; ANYAMBA, A.; SMALL, J.; PAK, E.; MASUOKA, E.; SCHMALTZ, J.; HANSEN, M.; PITTMAN, K.; BIRKETT, C.; WILLIAMS, D.; REYNOLDS, C.; DOORN, B. Monitoring Global Croplands with Coarse Resolution Earth Observations: The Global Agriculture Monitoring (GLAM) Project. **Remote Sensing**, v. 2, n. 6, p. 1589–1609, 2010.

BORGES, E. F.; SANO, E. E.; BORGES, E. F.; SANO, E. E. Séries Temporais de EVI do MODIS para o mapeamento de uso e cobertura vegetal do oeste da Bahia. **Boletim de Ciencias Geodesicas**, v. 20, n. 3, p. 526–547, 2014.

CARVALHO JÚNIOR, O. A.; SAMPAIO, C. da S.; SILVA, N. C.; COUTO JÚNIOR, A. F.; GOMES, R. A. T.; CARVALHO, A. P. F. de; SHIMABUKURO, Y. E. Classificação de padrões de savana usando assinaturas temporais NDVI do sensor MODLS no Parque Nacional Chapada dos Veadeiros. **Revista Brasileira de Geofísica**, v. 26, n. 4, p. 505–517, dez. 2008a.

CARVALHO JÚNIOR, O. A. de; COUTO JUNIOR, A. F.; SILVA, N. C. da; MARTINS, É. de S.; CARVALHO, A. P. F.; GOMES, R. A. T. Avaliação Dos Classificadores Espectrais De Mínima Distância Euclidiana E Spectral Correlation Mapper Em Séries Temporais Ndzi-Modis No Campo De Instrução Militar De Formosa (Go). **Revista Brasileira de Cartografia**, v. 4, n. 61, p. 399–412, 2009.

CARVALHO JÚNIOR, O. A. de; SAMPAIO, C. da S.; SILVA, N. C. da; COUTO JÚNIOR, A. F.; GOMES, R. A. T.; CARVALHO, A. P. F. de; SHIMABUKURO, Y. E. Classificação de padrões de savana usando assinaturas temporais NDVI do sensor MODLS no Parque Nacional Chapada dos Veadeiros. **Revista Brasileira de Geofísica**, v. 26, n. 4, p. 505–517, dez. 2008b.

CASTRO, C. N. de. **A Agropecuária Na Região Sudeste: Limitações E Desafios Futuros**. Brasília: IPEA, 2014.

CHANG, C.; LIN, C. LIBSVM: A Library for Support Vector Machines. **ACM Transactions on Intelligent Systems and Technology (TIST)**, v. 2, p. 1–39, 2013.

CHEN, C.-F.; CHEN, C.-R.; SON, N.-T. Investigating Rice Cropping Practices and Growing Areas from MODIS Data Using Empirical Mode Decomposition and Support Vector Machines. **GIScience & Remote Sensing**, v. 49, n. 1, p. 117–138, 2012.

CHEN, J.; JÖNSSON, P.; TAMURA, M.; GU, Z.; MATSUSHITA, B.; EKLUNDH, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. **Remote Sensing of Environment**, v. 91, n. 3–4, p. 332–344, 2004.

COHEN, W. B.; JUSTICE, C. O. Validating MODIS terrestrial ecology products: Linking in situ and satellite measurements. **Remote Sensing of Environment**, v. 70, n. 1, p. 1–3, 1999.

CONCEIÇÃO, J. C. P. R.; CONCEIÇÃO, P. H. Z. **Agricultura: Evolução e Importância para a Balança Comercial Brasileira**. Brasília - DF: IPEA, 2014.

CONGALTON, R. G. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. **Photogrammetric engineering and remote sensing**, v. 54, n. 5, p. 587–592, 1988a.

CONGALTON, R. G. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. **Photogrammetric Engineering and Remote Sensing**, v. 54, p. 593–600, 1988b. Disponível em: <<http://www.csa.com/partners/viewrecord.php?requester=gs&collection=TRD&recid=A8839098AH>>.

CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, v. 37, n. 1, p. 35–46, 1991.

CONGALTON, R. G.; GREEN, K. A practical look at the sources of confusion in error matrix generation. **Photogrammetric Engineering & Remote Sensing**, v. 59, n. 5, p. 641–644, 1993.

CONGALTON, R. G.; ODERWALD, R. G.; MEAD, R. A. Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques. **Photogrammetric Engineering and Remote Sensing**, v. 49, n. No. 12, p. 1671–1678, 1983.

COUTO JÚNIOR, A. F.; CARVALHO JÚNIOR, O. A.; MARTINS, É. D. S. Séries temporais MODIS aplicadas em sucessão de culturas de soja (*Glycine max* (L.) Merrill) e milho (*Zeamays* L.) em sistema de plantio direto. **Revista Brasileira de Cartografia**, v. 64, n. 3, p. 405–418, 2012a.

COUTO JÚNIOR, A. F.; CARVALHO JÚNIOR, O. A.; MARTINS, É. de S.; ENCINAS, J. I. Identificação das Assinaturas Temporais NDVI das Principais Fisionomias da Região do Ecomuseu do Cerrado. **Espaço & Geografia**, v. 10, n. 1, p. 173–189, 2007.

COUTO JÚNIOR, A. F.; CARVALHO JÚNIOR, O. A. de; MARTINS, É. D. S. Séries temporais MODIS aplicadas em sucessão de culturas de soja (*Glycine max* (L.) Merrill) e milho (*Zeamays* L.) em sistema de plantio direto. **Revista Brasileira de Cartografia**, v. 64, n. 3, p. 405–418, 2012b.

COUTO JUNIOR, A. F.; CARVALHO JUNIOR, O. A. de; MARTINS, E. de S.; SANTANA, O. A.; SOUZA, V. V. de; ENCINAS, J. I. Tratamento de ruídos e caracterização de fisionomias do Cerrado utilizando séries temporais do sensor MODIS. **Revista Árvore**, v. 35, n. 3 suppl 1, p. 699–705, 2011.

COVER, T.; HART, P. Nearest neighbor pattern classification. **IEEE Transactions on Information Theory**, v. 13, n. 1, p. 21–27, 1967. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1053964>>.

D'ARCO, E.; RUDORFF, B. R. T.; RIZZI, R. Índice de vegetação para estimativa de área plantada com arroz irrigado. In: RUDORFF, B. F. T.; SHIMABUKURO, Y. E.; CEBALLOS, J. C. (Ed.). **O sensor MODIS e suas aplicações ambientais no Brasil**. São José dos Campos - SP: Bookimage, 2007. p. 123–131.

DE CARVALHO JÚNIOR, O. A.; DA SILVA, N. C.; DE CARVALHO, A. P. F.; COUTO JÚNIOR, A. F.; SILVA, C. R.; SHIMABUKURO, Y. E.; GUIMARÃES, R. F.; GOMES, R. A. T. Combining Noise-Adjusted Principal Components transform and median filter techniques for denoising MODIS temporal signatures. **Revista Brasileira de Geofísica**, v. 30, n. 2, p. 147–157, 2012.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA). **Pesquisa, desenvolvimento e inovação para o agronegócio brasileiro: Cenários 2002-2012**. Brasília - DF: Embrapa, Secretaria de Gestão e Estratégia, 2003.

FISCHER, A. A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. **Remote Sensing of Environment**, v. 48, n. 2, p. 220–230, 1994.

FOODY, G. M. Status of land cover classification accuracy assessment. **Remote Sensing of Environment**, v. 80, n. 1, p. 185–201, 2002.

FOODY, G. M. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. **Photogrammetric Engineering & Remote Sensing**, v. 70, n. 5, p. 627–633, 2004.

FORKEL, M.; CARVALHAIS, N.; VERBESSELT, J.; MAHECHA, M. D.; NEIGH, C. S. R.; REICHSTEIN, M. Trend Change detection in NDVI time series: Effects of inter-annual variability and methodology. **Remote Sensing**, v. 5, n. 5, p. 2113–2144, 2013.

FRANCO-LOPEZ, H.; EK, A. R.; BAUER, M. E. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. **Remote Sensing of Environment**, v. 77, n. 3, p. 251–274, 2001.

GALFORD, G. L.; MUSTARD, J. F.; MELILLO, J.; GENDRIN, A.; CERRI, C. C.; CERRI, C. E. P. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. **Remote Sensing of Environment**, v. 112, n. 2, p. 576–587, 2008.

GALLO, K. P.; FLESCHE, T. K. Large-area crop monitoring with the NOAA AVHRR: Estimating the silking stage of corn development. **Remote Sensing of Environment**, v. 27, n. 1, p. 73–80, 1989.

GAO, T.; ZHU, J.; ZHENG, X.; SHANG, G.; HUANG, L.; WU, S. Mapping spatial distribution of larch plantations from multi-seasonal landsat-8 OLI imagery and multi-scale textures using random forests. **Remote Sensing**, v. 7, n. 2, p. 1702–1720, 2015.

GITELSON, A. A.; KAUFMAN, Y. J.; MERZLYAK, M. N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. **Remote Sensing of Environment**, v. 58, n. 3, p. 289–298, 1996.

GRAPS, A. An introduction to wavelets. **IEEE Computational Science and Engineering**, v. 2, n. 2, p. 50–61, 1995.

GUSSO, A.; FORMAGGIO, A. R.; RIZZI, R.; ADAMI, M.; RUDORFF, B. F. T. Soybean crop area estimation by Modis/Evi data. **Pesquisa Agropecuaria Brasileira**, v. 47, n. 3, p. 425–435, 2012.

HARTFIELD, K. A.; LANDAU, K. I.; VAN LEEUWEN, W. J. D. Fusion of high resolution aerial multispectral and lidar data: Land cover in the context of urban mosquito habitat. **Remote Sensing**, v. 3, n. 11, p. 2364–2383, 2011.

HAYKIN, S. **Redes Neurais, Princípios e Prática**. Porto Alegre: Bookman, 2001.

HELFAND, S.; REZENDE, G. **Padrões regionais de crescimento da produção de grãos e o papel da região Centro-Oeste**. Brasília - DF: IPEA, 2000.

HENDERSON, K. E.; BADHWAR, G. D. An initial model for estimating soybean development stages from spectral data. **Remote Sensing of Environment**, v. 14, n. 1–3, p. 55–63, 1984.

HUETE, A.; DIDAN, K.; MIURA, T.; RODRIGUEZ, E. .; GAO, X.; FERREIRA, L. . Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**, v. 83, n. 1–2, p. 195–213, nov. 2002.

[CSL STYLE ERROR: reference with no printed form.]

HUETE, A. R.; LIU, H. Q.; BATCHILY, K.; J., L. van W. A comparsion of vegetation indices over a Global set of TM images for EO -MODIS. **Remote Sensing of Environment**, v. 59, p. 440–451, 1997.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Levantamento Sistemático da Produção Agrícola**. Rio de Janeiro: IBGE, 2015.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Levantamento Sistemático Da Produção Agrícola**. Rio de Janeiro: IBGE, 2016. v. 29

JAKUBAUSKAS, M. E.; LEGATES, D. R.; KASTENS, J. H. Crop identification using harmonic analysis of time-series AVHRR NDVI data. **Computers and Electronics in Agriculture**, v. 37, n. 1–3, p. 127–139, 2002.

JIANG, Z.; HUETE, A. R.; DIDAN, K.; MIURA, T. Development of a two-band enhanced vegetation index without a blue band. **Remote Sensing of Environment**, v. 112, n. 10, p. 3833–3845, 2008.

JÖNSSON, P.; EKLUNDH, L. Seasonality extraction by function fitting to time-series of satellite sensor data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 40, n. 8, p. 1824–1832, 2002.

JÖNSSON, P.; EKLUNDH, L. TIMESAT - A program for analyzing time-series of satellite sensor data. **Computers and Geosciences**, v. 30, n. 8, p. 833–845, 2004.

JULIEN, Y.; SOBRINO, J. A. Global land surface phenology trends from GIMMS database. **International Journal of Remote Sensing**, v. 30, n. 13, p. 3495–3513, 2009.

JUSTICE, C.; BECKER-RESHEF, I. **Report from the Workshop on Developing a Strategy for Global Agricultural Monitoring in the framework of Group on Earth Observations (GEO) 16-18 July 2007, FAO, Rome**. [s.l: s.n.].

JUSTICE, C. O.; TOWNSHEND, J. R. G.; VERMOTE, E. F.; MASUOKA, E.; WOLFE, R. E.; SALEOUS, N.; ROY, D. P.; MORISETTE, J. T. An overview of MODIS Land data processing and product status. **Remote Sensing of Environment**, v. 83, n. 1–2, p. 3–15, 2002.

JUSTICE, C. O.; VERMOTE, E. F.; TOWNSHEND, J. R. G.; DEFRIES, R. S.; ROY, D. P.; HALL, D. K.; SALOMONSON, V. V; PRIVETTE, J. L.; RIGGS, G.; STRAHLER, A. H.; LUCHT, W.; MYNENI, R. B.; KNYAZIKHIN, Y.; RUNNING, S. W.; NEMANI, R. R.; WAN, Z.; HUETE, A. R.; VAN LEEUWEN, W.; WOLFE, R. E.; GIGLIO, L.; MULLER, J. P.; LEWIS, P.; BARNESLEY, M. The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. **Geoscience and Remote Sensing, IEEE Transactions on**, v. 36, n. 4, p. 1228–1249, 1998. Disponível em:

<http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=701075%5Cnpapers2://publication/doi/10.1109/36.701075>.

KRUSE, F. A.; LEFKOFF, A. B.; BOARDMAN, J. W.; HEIDEBRECHT, K. B.; SHAPIRO, A. T.; BARLOON, P. J.; GOETZ, A. F. H. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. **Remote Sensing of Environment**, v. 44, n. 2–3, p. 145–163, 1993.

KUPLICH, T. M.; MOREIRA, A.; FONTANA, D. C. Série temporal de índice de vegetação sobre diferentes tipologias vegetais no Rio Grande do Sul. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 17, n. 55, p. 1116–1123, 2013.

LAMPARELLI, R. A. C.; CARVALHO, W. M. O. de; MERCANTE, E. Mapeamento de semeaduras de soja (*Glycine max* (L.)Merr.) mediante dados MODIS/Terra E TM/Landsat 5: um comparativo. **Engenharia Agrícola**, v. 28, n. 2, p. 334–344, 2008.

LEEuw, J. D.; JIA, H.; YANG, L.; SCHMIDT, K.; SKIDMORE, A. K. Comparing accuracy assessment to infer superiority of image classification methods. **International Journal of Remote Sensing**, v. 27, n. 1, p. 223–232, 2006.

LOHMANN, M.; DEPPE, F.; SIMÕES, K.; MARTINI, L. Monitoramento da evolução temporal da cultura da soja no estado do Paraná utilizando imagens do sensor MODIS. In: XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal-RN. Natal-RN: 2009.

MARQUEZOLA, D. S.; SANTOS, J. W. M. C. Mapeamento do uso da Terra e Desenvolvimento da Cultura da Soja Numa Fazenda do Sul de Mato Grosso a partir de Dados Medidos in loco e de Imagens Orbitais. In: XVI Simpósio Brasileiro de Sensoriamento Remoto, Foz do Iguaçu-PR. Foz do Iguaçu-PR: 2013.

MINGWEI, Z.; QINGBO, Z.; ZHONGXIN, C.; JIA, L.; YONG, Z.; CHONGFA, C. Crop discrimination in Northern China with double cropping systems using Fourier analysis of time-series MODIS data. **International Journal of Applied Earth Observation and Geoinformation**, v. 10, n. 4, p. 476–485, 2008.

MONTIBELLER FILHO, G.; GARGIONI, S. L. Desenvolvimento da Região Sul do Brasil. In: MONTORO, G. C. F.; GUIDOLIN, S. M. (Ed.). **Desenvolvimento da região sul do Brasil**. Rio de Janeiro: BNDES, 2014. p. 310–325.

MOULIN, S.; KERGOAT, L.; VIOVY, N.; DEDIEU, G. Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements. **Journal of Climate**, v. 10, n. 6, p. 1154–1170, 1997.

OZDOGAN, M. The spatial distribution of crop types from MODIS data: Temporal unmixing using Independent Component Analysis. **Remote Sensing of Environment**, v. 114, n. 6, p. 1190–1204, 2010.

PAN, Y.; LI, L.; ZHANG, J.; LIANG, S.; ZHU, X.; SULLA-MENASHE, D. Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. **Remote Sensing of Environment**, v. 119, p. 232–242, 2012.

PANDA, S. S.; AMES, D. P.; PANIGRAHI, S. Application of vegetation indices for agricultural crop yield prediction using neural network techniques. **Remote Sensing**, v. 2, n. 3, p. 673–696, 2010.

PENG, Y.; GITELSON, A. A.; SAKAMOTO, T. Remote estimation of gross primary productivity in crops using MODIS 250m data. **Remote Sensing of Environment**, v. 128, p. 186–196, 2013.

PIFFER, T. R. O.; AMARAL, G. M.; CAMPOS, P. M.; SILVA, R. A. B. Mapeamento de áreas cultivadas. In: **Agricultura e Abastecimento Alimentar: políticas públicas e mercado agrícola**. Brasília: Conab, 2009. p. 312–322.

PITTMAN, K.; HANSEN, M. C.; BECKER-RESHEF, I.; POTAPOV, P. V.; JUSTICE, C. O. Estimating global cropland extent with multi-year MODIS data. **Remote Sensing**, v. 2, n. 7, p. 1844–1863, 2010.

REZENDE, G. **Ocupação agrícola e estrutura agrária no Cerrado: o papel dos recursos naturais e da tecnologia**. Brasília: IPEA, 2002.

RICHARDS, J. a. **Remote Sensing Digital Image Analysis: An Introduction**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999.

RIZZI, R.; RUDORFF, B. F. T. Estimativa da área de soja no Rio Grande do Sul por meio de imagens landsat. **Revista Brasileira de Cartografia**, v. 57, n. 3, p. 226–234, 2005.

ROBERTSON, L. D.; KING, D. J. Comparison of pixel-and object-based classification in land cover change mapping. **International Journal of Remote Sensing**, v. 32, n. 6, p. 1505–1529, 2011.

ROSA, V. G. C. da; MOREIRA, M. A.; RUDORFF, B. F. T.; ADAMI, M. Estimativa da produtividade de café com base em um modelo agrometeorológico-espectral. **Pesquisa Agropecuária Brasileira**, v. 45, n. 12, p. 1478–1488, 2010.

ROSEMBACK, R.; FERREIRA, N. J.; SHIMABUKURO, Y. E.; CONFORTE, J. C. Análise de ondaletas aplicada a dados multitemporais de NDVI/MODIS. **Geografia**, v. 34, n. 3, p. 559–575, 2009.

SAKAMOTO, T.; WARDLOW, B. D.; GITELSON, A. A.; VERMA, S. B.; SUYKER, A. E.; ARKEBAUER, T. J. A Two-Step Filtering approach for detecting maize and soybean phenology with time-series MODIS data. **Remote Sensing of Environment**, v. 114, n. 10, p. 2146–2159, 2010.

SAKAMOTO, T.; YOKOZAWA, M.; TORITANI, H.; SHIBAYAMA, M.; ISHITSUKA, N.; OHNO, H. A crop phenology detection method using time-series MODIS data. **Remote Sensing of Environment**, v. 96, n. 3–4, p. 366–374, 2005.

SANTANA, O. A.; JÚNIOR, O. A. de C.; PIMENTEL, C. M. M.; GOMES, R. A. T.; OLIVEIRA, S. N. Modelagem de espectros temporais NDVI-MODIS, no período de 2000 a 2008, na bacia do Rio Paracatu, Brasil. **Revista Brasileira de Geofísica**, v. 28, n. 1, p. 47–60, 2010.

SANTOS, E. C.; SOUZA, A. L. F.; AMARAL, G. M.; SOUZA, A. O.; PIFFER, T. R. O.; SOUZA, L. M. M.; OLIVEIRA, C. C. Mapeamento da cultura do café na microrregião de Afonso Cláudio – Espírito Santo, com imagens de aerofotogrametria. In: XV Simpósio Brasileiro de Sensoriamento Remoto, Curitiba - PR. Curitiba - PR: 2011.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627–1639, 1964.

SILVA, M. A. D. A.; ZEILHOFER, P.; SCHWENK, L. M.; FEDERAL, U.; GROSSO, D. M.; GEOGRAFIA, M. Classificação Digital De Imagens Modis Para Mapeamento Do Uso Da Terra Na Bacia Do Médio E Alto Teles Pires-Mt . p. 1–6, 2010.

SOUSA, A. O.; SOUZA, A. L. F.; PIFFER, T. R. O.; AMARAL, G. M.; SANTOS, E. C.; SOUZA, L. M. M.; OLIVEIRA, C. C. Geoprocessamento e sensoriamento como subsídio à previsão de safra de café no Extremo Oeste Baiano. In: XV Simpósio Brasileiro de Sensoriamento Remoto, Curitiba - PR. Curitiba - PR: 2011.

SOUZA, L. M. M.; SOUZA, A. L. F.; AMARAL, G. M.; PIFFER, T. R. O.; SOUSA, A. O.; SANTOS, E. C.; OLIVEIRA, C. C. Sensoriamento Remoto e geoprocessamento aplicados ao mapeamento das lavouras de café nas microrregiões de Ji-Paraná, Vilhena e Colorado Do Oeste em Rondônia. In: XV Simpósio Brasileiro de Sensoriamento Remoto, Curitiba - PR. Curitiba - PR: 2011.

STEHRMAN, S. V. Selecting and interpreting measures of thematic classification accuracy. **Remote Sensing of Environment**, v. 62, n. 1, p. 77–89, 1997.

STEHRMAN, S. V. Estimating the Kappa coefficient and its variance under stratified random sampling. **Photogrammetric Engineering & Remote Sensing**, v. 62, n. 4, p. 401–407, 1996.

TUCKER, C. J.; ELGIN, J. H.; MCMURTREY, J. E.; FAN, C. J. Monitoring corn and soybean crop development with hand-held radiometer spectral data. **Remote Sensing of Environment**, v. 8, n. 3, p. 237–248, 1979.

ULLMANN, T.; SCHMITT, A.; ROTH, A.; DUFFE, J.; DECH, S.; HUBBERTEN, H. W.; BAUMHAUER, R. Land cover characterization and classification of arctic tundra environments by means of polarized synthetic Aperture X- and C-band radar (PolSAR) and Landsat 8 multispectral imagery - Richards Island, Canada. **Remote Sensing**, v. 6, n. 9, p. 8565–8593, 2014.

VAN OORT, P. A. J. Interpreting the change detection error matrix. **Remote Sensing of Environment**, v. 108, n. 1, p. 1–8, 2007.

VERMOTE, E. F.; EL SALEOUS, N. Z.; JUSTICE, C. O. Atmospheric correction of MODIS data in the visible to middle infrared: First results. **Remote Sensing of Environment**, v. 83, n. 1–2, p. 97–111, 2002.

VICTORIA, D. de C.; DA PAZ, A. R.; COUTINHO, A. C.; KASTENS, J.; BROWN, J. C. Cropland area estimates using Modis NDVI time series in the state of Mato Grosso, Brazil. **Pesquisa Agropecuaria Brasileira**, v. 47, n. 9, p. 1270–1278, 2012.

WAGENSEIL, H.; SAMIMI, C. Assessing spatio-temporal variations in plant phenology using Fourier analysis on NDVI time series: Results from a dry savannah environment in Namibia. **International Journal of Remote Sensing**, v. 27, n. 16, p. 3455–3471, 2006.

WANG, Q.; TENHUNEN, J. D. Vegetation mapping with multitemporal NDVI in North Eastern China Transect (NECT). **International Journal of Applied Earth Observation and Geoinformation**, v. 6, n. 1, p. 17–31, 2004.

WHITCRAFT, A. K.; BECKER-RESHEF, I.; JUSTICE, C. O. Agricultural growing season calendars derived from MODIS surface reflectance. **International Journal of Digital Earth**, v. 8, n. 3, p. 173–197, 2015.

XU, Y.; SHI, J.; DU, J. An improved endmember selection method based on vector length for MODIS reflectance channels. **Remote Sensing**, v. 7, n. 5, p. 6280–6295, 2015.

ZHANG, X.; FRIEDL, M. A.; SCHAAF, C. B.; STRAHLER, A. H.; HODGES, J. C. F.; GAO, F.; REED, B. C.; HUETE, A. Monitoring vegetation phenology using MODIS. **Remote Sensing of Environment**, v. 84, n. 3, p. 471–475, 2003.

ZHANG, X.; HODGES, J. C. F.; SCHAAF, C. B.; FRIEDL, M. A.; STRAHLER, A. H.; GAO, F. Global Vegetation Phenology from AVHRR and MODIS Data. **Geoscience and Remote Sensing Symposium**, p. 2262–2264, 2001.

CAPÍTULO 2 – COMPARING BACK-PROPAGATION NEURAL NETWORKS AND SUPPORT VECTOR MACHINE FOR CROPPING SYSTEM MAPPING USING MODIS NDVI TIME SERIES IN MATO GROSSO STATE, BRAZIL

Abstract: This research evaluates the Neural Network and Support Vector Machine methods to identify the natural vegetation and crop types in six municipalities of the agricultural frontier of Mesoregion North of Mato Grosso state through the Moderate Resolution Imaging Spectrometer (MODIS) Normalized Difference Vegetation Index (NDVI) product. The image processing consisted of the following steps: (1) acquisition of MODIS images, (2) removal of noise, (3) time-series classification using Artificial Neural Network (ANN) and Support Vector Machine (SVM) considering different kernel functions (linear, polynomial, RBF and Sigmoid), and (4) accuracy analysis using confusion matrix, Overall Accuracy, Kappa index and McNemar's test. The NDVI temporal signatures allowed us to identify the seasonal dynamics of following classes: forest formation, savanna, pasture, annual single cropping system (soybean and corn), double cropping system (soybean/corn and soybean/cotton), and center pivot (triple cropping system). The network performance in crop mapping considered different geometry and internal parameters. The best ANN result with Kappa index of 0.77 adopted one hidden layer, training threshold contribution of 0.7, training rate of 0.2, and training momentum of 0.4. The McNemar's test compared the ANN models with the best Kappa indices, identifying statistically significant changes in the classification results that reject the null hypothesis of no processing effect. In contrast, the McNemar's test in the SVM models demonstrates that the methods using linear, polynomial and RBF functions had equivalent results (except the sigmoid function), despite the slight variation of Kappa index (0.73-0.75) and overall accuracy (77.57-78.9). Moreover, the results of these SVM models compared to the best ANN model are statistically equivalent to the McNemar's test, demonstrating the efficiency of both methods in the classification of agricultural planting areas. The SVM method is a preferred choice over the ANN method due to its high quality of results considering the different kernel functions and does not require an extensive evaluation of the initial parameters. Therefore, time series analysis of remotely sensed data provides significant progress in the information acquisition about the spatial distribution of crop in an important Brazilian agricultural region.

Keywords: Time-series analysis; Vegetation phenology; Neural Network; Support Vector Machine, Image Processing, Cropland Identification, Noise Removal.

2.1. INTRODUCTION

Brazil has about 57.3 million hectares of agricultural land and 199.7 million tons of grain production (IBGE, 2015). The West-Central region is the leading grain producer of the country containing 40.3% of national production, while the Mato Grosso State leads the production with a share of 23.6% (IBGE, 2015). Development of agriculture in the West-Central region has advantages over other Brazilian regions, because of the suitable climate, good quality of soil physical properties, vast arable land and flat topography adequate to

mechanization (CASTRO, 2014). Moreover, the advent of technological innovations provided the agricultural growth in this region such as the transformation of low fertility and high acidity soil to fertile soils, new seed varieties, genetic improvement and adaptations of agricultural management (REZENDE, 2002).

Agricultural policies rely on information on food production to develop strategies and mechanisms that control farm stock and food prices. However, agricultural crop confronts interference of different factors such as weather (flooding, drought, and frost), socioeconomic, energy and political, which affect the guarantee of food security. In this context of the global market, agricultural monitoring systems play a fundamental role in decision-making processes and regional development planning. The main international monitoring systems include Global Information and Early Warning System (GIEWS) developed by the Food and Agriculture Organization (FAO), Monitoring Agriculture by Remote Sensing (MARS) of the European Union, Global Agriculture Monitoring (GLAM) of the United States, and Crop Watch System (CAS) of China (BECKER-RESHEF et al., 2010). Monitoring and forecasting of agricultural production in Brazil is a significant challenge because of regional differences in soil, topography, climate, management practices, diseases and agrarian calendar. Currently, an estimated cultivated area of the main grain crops in Brazil adopts data acquired by interviews together farmers by National Food Supply Agency (CONAB) and Brazilian Institute of Geography and Statistics (IBGE).

Remote sensing techniques have become a useful tool for monitoring agricultural areas and crop forecasting, enabling information about the Earth's surface with less cost and time (KASTENS et al., 2005; ATZBERGER, 2013). The spectral behavior of cultures changes according to the phenological stages, photosynthetic activity, and spacing between plants. Individual analysis of a single image throughout the year may be difficult to characterize the type of agricultural crop. However, remote sensing images with high temporal resolution bring new perspectives to solve the problem by allowing describe and distinguish different crop phenologies. In this approach, different digital processing techniques using remotely sensed time-series data had been developed for agricultural monitoring (SAKAMOTO et al., 2005; LUNETTA et al., 2006; PRINGLE; DENHAM; DEVADAS, 2012).

The Moderate Resolution Imaging Spectroradiometer (MODIS) sensor is widely used in the temporal analysis because of its high temporal resolution (almost daily repetition) and moderate spatial resolution (250m and 500m). The MODIS time series had been widely used both in mapping of natural areas (ZHANG et al., 2003; COUTO JUNIOR et al., 2011; CARVALHO JÚNIOR et al., 2012), and agricultural areas such as rice (XIAO et al., 2005, 2006, GUMMA et al., 2011, 2015; TELUGUNTLA et al., 2015), sugarcane (XAVIER et al., 2006; AGUIAR et al., 2008; ANTUNES; LAMPARELLI; RODRIGUES, 2015), coffee (ROSA et al., 2010; COUTO JÚNIOR et al., 2013), wheat (OZDOGAN, 2010; PAN et al., 2012; FRANCH et al., 2015), soybean (SAKAMOTO et al., 2010; BERNARDES et al., 2011; COUTO JÚNIOR; CARVALHO JÚNIOR; MARTINS, 2012), and corn (SAKAMOTO; WARDLOW; GITELSON, 2011; SAKAMOTO; GITELSON; ARKEBAUER, 2013). The use of time series in Brazil still is limited in agricultural monitoring because this country has a

variety of different crops and growth behavior. Thus, the aim of this research is to compare SVM and back-propagation neural network methods of MODIS time-series classification considering the producing regions of soybeans, cotton, and corn in Central Brazil (Mesoregion Northern Mato Grosso State) during the 2013/14 growing season.

2.2. MATERIALS AND METHODS

2.2.1. Study Area

The study area has 32,915 square kilometers located in North Mesoregion of Mato Grosso state, comprising the following municipalities: Sinop, Ipiranga do Norte, Sorriso, Vera, Lucas do Rio Verde and Nova Mutum (**Figure 2.1**). This region has a tropical continental climate with rainfall totals between 1800-2100 mm and an average annual temperature above 18°C (CASTRO JÚNIOR; PAULO, 2002). The climate has two distinct seasons: a dry season from May to September (250-300 mm) and a rainy season from November to April (800-1100 mm).

The study area is on the Parecis Plateau with altitudes ranging from 350 to 550 m composed of conglomerates and sandstones, fluvial and aeolian deposits of the Late Cretaceous (BAHIA et al., 2006; BAHIA; NETO, 2007). The rocks in the region belong to the Parecis Group, which divided into two distinct sedimentary packages: Salto das Nuvens Formation in the basal portion and Utariti Formation at the top (BARROS et al., 1982). Salto das Nuvens Formation consists of conglomerates intercalated with red sandstone lenses and bimodal sandstones, which represented sedimentary deposits in alluvial fan environments and river channel, containing aeolian system contribution (BARROS et al., 1982; WESKA, 2006). The Utariti formation has sandstones with several colors, formed mainly of quartz and feldspar, containing cross-stratified and plane-parallel laminated sand beds (BARROS et al., 1982; CUTRIM et al., 2007). The vegetation consists of Savannah physiognomies of the Cerrado biome.

Environmental factors are favorable for agriculture, where the six municipalities have 1,586,550 hectares of soybean planted, 651,120 hectares of corn with a significant cotton production (IBGE, 2015). The municipality of Sorriso is the largest producer of soybeans in Mato Grosso state corresponding to 9.30% of the state production, being known as the greatest producer of soybeans in Brazil (IBGE, 2015).

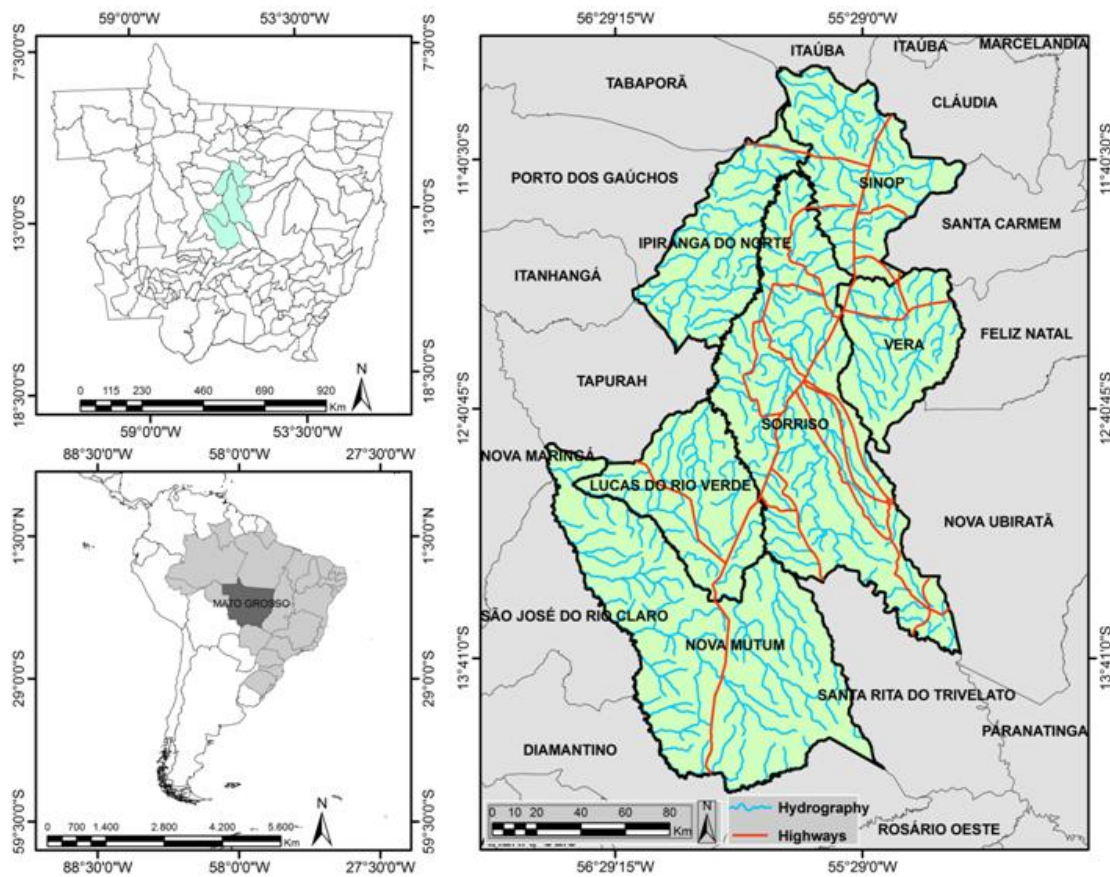


Figure 2.1. Location of the study area.

2.2.2. MODIS Data

We used the MODIS Surface Reflectance product (MOD09Q1) with 250-meter resolution, which includes red (620–670 nm) and near-infrared (841–875 nm) bands. This 8-day composite product selects pixels of the highest quality based on a combination of scenes to reduce noise and atmospheric effects (aerosols, clouds, and cloud shadows). The high temporal resolution MODIS images promote an increase in the cloudless data acquisition, allowing a continuous monitoring of the crops. In this research, MODIS data correspond to the tile h12v10 acquired on the website of the United States Geological Survey (<http://earthexplorer.usgs.gov/>). Original images in sinusoidal projection were re-projected to the geographic coordinate system (WGS 84 ellipsoid), using the MODIS Reprojection Tool (MRT) software (DWYER; SCHMIDT, 2006). The analysis period included 46 images between October 2013 and October 2014, which correspond to start and end date of the crop cycles in Mato Grosso State.

Normalized Difference Vegetation Index (NDVI) was calculated from the surface reflectance values of red (pRED) and near infrared (pNIR) (ROUSE et al., 1973). The MODIS Vegetation Index Products (MOD13Q1 and MOD13A1) were not used because of the temporal resolution (16-day composite) lower than the MOD09 (8-day composite), which are not appropriate to describe the savanna and crop phenologies (ABADE et al., 2015). In

multiple cropping systems, the frequency of high and low peaks in time series may be small not being evidenced in the 16-day composite images.

2.2.3. NDVI Time Series Noise Reduction

The noises are common in NDVI time series affecting the identification of surface targets from the cloud cover, shadows, and instrumental defects. Different noise-reduction methods for time-series of satellite sensor data have been proposed (HIRD; MCDERMID, 2009; ATKINSON et al., 2012; GENG et al., 2014). Several strategies for constructing a high-quality NDVI time-series are based on two steps: (a) removal of spikes and outliers within the time-series data, and (b) data smoothing using a filter (JÖNSSON; EKLUNDH, 2002; CHEN et al., 2004). In these two steps, we used the Savitzky and Golay (S-G)'s method (CARVALHO JÚNIOR et al., 2015). The S-G filter was originally developed for chemical spectrum (SAVITZKY; GOLAY, 1964), but it demonstrates efficiency to remove noise contamination in the MODIS time series (CHEN et al., 2004). The S-G filter combines the data smoothing and the waveform-peak preservation (height, shape, and asymmetry), which is significant for the phenological studies from satellite image time-series. Several surveys confirm the S-G filter efficiency in coupling the noise elimination and the preservation of phenological attributes (REN et al., 2008; HÜTTICH et al., 2009; GENG et al., 2014).

The preprocessing step to remove spikes and outlier considers the highest values of the difference between the original and filtered data (CARVALHO JÚNIOR et al., 2015; SHAO et al., 2016). In this paper, the method for detecting outliers calculates the absolute value of subtraction between original data and S-G curve from window size of nine pixels. If the absolute value exceeds a threshold value determined by the user, it is considered as a spike/outlier and is replaced by the linear interpolation value between the points before and after the outliers. In the second step, we used the S-G method for smoothing the time series, considering a window size of nine pixels.

2.2.4. MODIS Time-Series Classification

We compared two supervised classification methods: support vector machines (SVMs) and neural network classification using standard back-propagation. The artificial neural network (ANN) has been widely applied in the classification of remotely sensed data (MAS; FLORES, 2008). Recently, ANN methods have been tested to map the spatial distribution of crops types from time series data (LUNETTA et al., 2010; ATZBERGER; REMBOLD, 2013). In this research, we used the Neural Net from ENVI software, which uses standard back-propagation and logistic activation function for supervised learning (RICHARDS, 2013). The back-propagation method is the most commonly used and best-known training algorithm for ANNs, presenting versatility, robustness, and high precision enabling nonlinear classifications. The choice of internal parameters of the back-propagation network is one of the most complicated factors, usually determined using trial and error (MAIER; DANDY, 1998). The Neural Net of the ENVI program has the following input parameters: training threshold contribution, training rate, training momentum, training RMS exit criteria, number of hidden layers, and number of training iterations. The combinations of these factors result in variable performances, making the internal comparison of this method a separate analysis in

relation to the other classification procedures. A set of empirical tests were performed to evaluate the back-propagation neural networks efficiency, considering the following values for the internal parameters: training threshold contribution (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9); training rate (0.005, 0.02, 0.1, 0.2, 0.4, 0.6, 0.8); training momentum (0.05, 0.2, 0.4, 0.6, 0.8, and 0.9); training RMS exit criteria (0.1); and number of training iterations (5,000). We investigated two network geometries, adopting single and two hidden layers. The different parameters tested in the best-fit model identification are shown for the case of crop mapping.

SVMs consist of a method set for supervised non-parametric statistical learning, which establishes a perceptual representation model (SHMILOVICI, 2005). The mapping of the training vectors obtains an optimal separation from the higher-dimensional dataset into discrete classes, considering the hyperplane description from a kernel function (SCHÖLKOPF, 2002). Therefore, SVM learning process does not use prior assumptions of the probability distribution. Mountrakis et al. (MOUNTRAKIS; IM; OGOLE, 2011) review the utilization of the support vector machine (SVMs) for remote sensing data, showing the exponentially growing number of recent research. We used the SVM with four kernel functions: linear, polynomial, radial basis function (RBF), and sigmoid.

In the different classification methods, the land-use/land-cover classes were forest, savanna, pasture, annual single cropping system, soybean/corn and soybean/cotton double cropping systems, and center pivot (triple cropping system). Three annual single cropping systems occur in the study area (soybean, cotton, and corn), which were classified independently and then merged into a single class because of their low representativeness. We use a mask isolating water bodies and urban areas to eliminate possible misclassifications in those localities.

2.2.5. Accuracy Analysis

The accuracy analysis is crucial both to establish the product quality and to compare classification algorithms. We used the confusion matrix, Kappa coefficient, Overall accuracy, and commission and omission errors for the accuracy assessment of land-use/land-cover classifications (CONGALTON; ODERWALD; MEAD, 1983). The accuracy of the different methodological approaches adopted 300 validation sites for each class (ground-truth points) collected from a systematic, stratified sampling design (FOODY, 2002). The selection of sampling points considered fieldwork and visual interpretation of Landsat images 8 to produce a spatially well-distributed sample.

Additionally, we compared the classification methods using the McNemar's test to detect statistically significant differences between the results (MCNEMAR, 1947). This procedure is widely used in remote sensing studies, where related samples are used to compare the accuracy of the thematic maps (FOODY, 2004). McNemar's test assesses the significance of marginal proportions from a 2x2 contingency table considering a binary distinction of correct and incorrect points for two classification results (**Table 2.1**).

Table 2. 1 - Data layout for McNemar test between two classification results

	Classification 2		
	Correct	Incorrect	Total
Classification 1			
Correct	f_{11}	f_{12}	$f_{11}+f_{12}$
Incorrect	f_{21}	f_{22}	$f_{21}+f_{22}$
Total	$f_{11}+f_{21}$	$f_{12}+f_{22}$	$f_{11}+f_{12}+f_{21}+f_{22}$

In this test, the chi-square (χ^2) statistic has one degree of freedom and uses exclusively discordant samples (f_{12} e f_{21}) (equation 1). The null hypothesis of marginal homogeneity means equivalent results between as two classifications. If the result χ^2 is greater than the values of the χ^2 distribution table, the null hypothesis is rejected by assuming that the marginal proportion of each classification method is significant and with different behavior. In this research, we adopted 2100 samples in the McNemar's test, which correspond to the same set of samples applied in the accuracy analysis.

$$\frac{(f_{12} - f_{21})^2}{f_{12} + f_{21}} \quad (\text{eq.1})$$

2.3. RESULTS

2.3.1. Temporal signatures

NDVI temporal signatures of the classes in the study area for the 2013/2014 crop year were evaluated. The crop time series presented a variation of photosynthetic intensity throughout the year (**Figure 2.2**). A single peak in the temporal signature characterized the unique annual planting system. The cotton crop had minimum NDVI values (0.3) in February and maximum (0.8) in June, decreased in early July. The corn crop had minimum NDVI values (0.3) in March and maximum (0.8) in May, falling in July. Finally, the soybean plantation had the minimum NDVI values (0.2) in October and the maximum (0.7) in February, decreasing in March (**Figure 2.2a**).

The double annual planting system (soybean/cotton and soybean/corn) had two peaks in the temporal signatures, which the first represented soybean and the second the cotton or the corn (**Figure 2.2b**). The cotton-peak was wider than the corn peak. In succession soybean/cotton, the soybean curve had the minimum NDVI value (0.2) in October, the maximum value (0.7) in January and return to decline in February, while cotton curve had the minimum NDVI value (0.2) in March, the maximum (0.8) in June and a decrease in July. In succession soybean/corn, the soybean curve had the same characteristics previously described, and the corn curve had the minimum value (0.2) in March, the maximum value (0.8) in May and a subsequent fall in June.

Normally, the center pivots had triple annual planting system characterized by three peaks in NDVI time series (**Figure 2.2c**). These peaks occurred in January (0.7), March (0.7), and September (0.7). The time series of forest formation had a different behavior of the

cultivation areas, comprising the lowest and highest NDVI values (**Figure 2.2a**). The time series of pasture showed a minimum NDVI value of 0.4 and maximum of 0.7. The savanna described minimum NDVI value of 0.5 and maximum of 0.68.

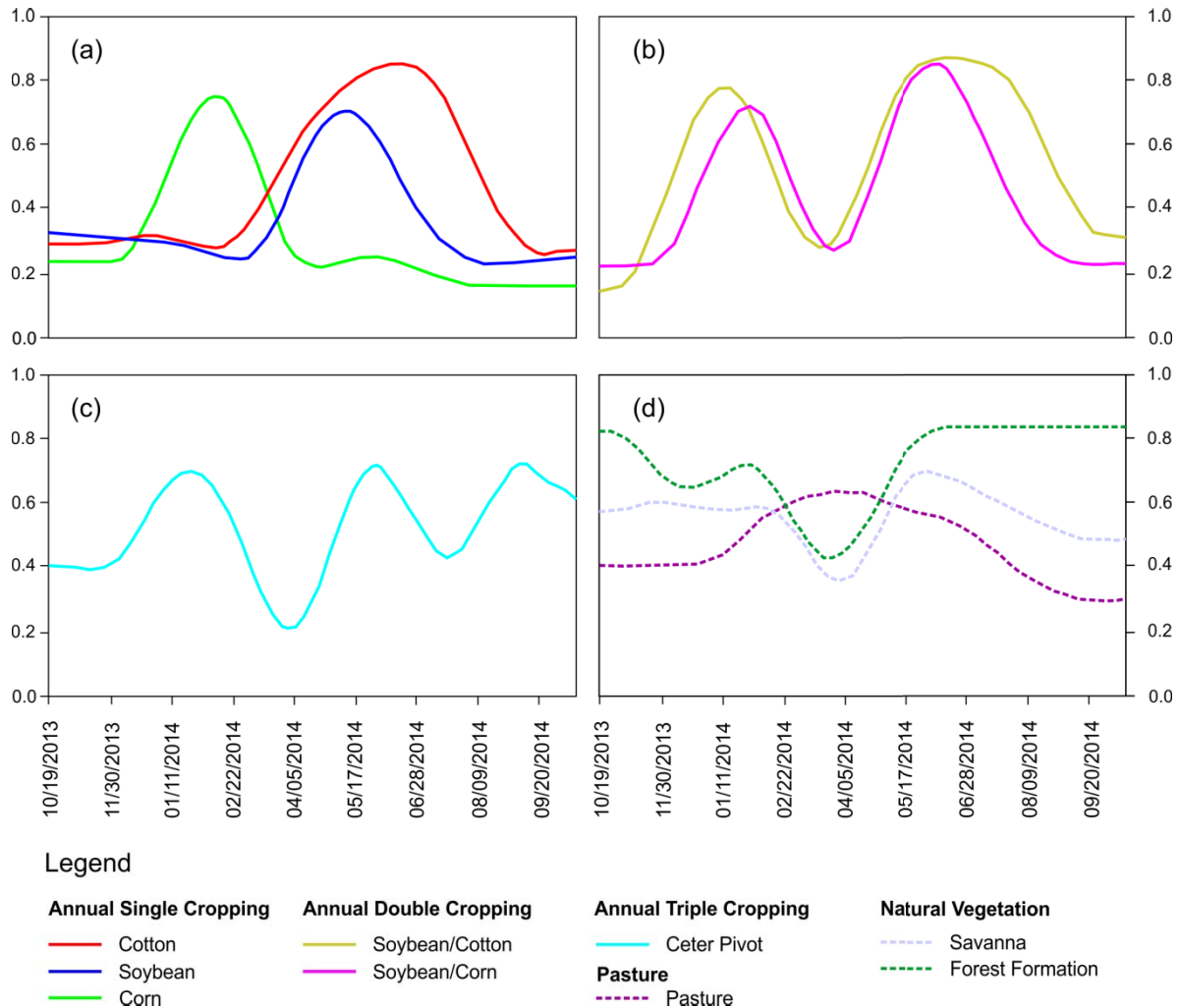


Figure 2.2. Temporal signatures of land use and crop classes: (a), (b), (c) and (d).

2.3.2. Results of time-series classifications using ANN

Only the ANN models with single hidden layer obtained good results in the tests. The ANN models with two hidden layers performed unsatisfactorily, producing a broad generalization of classes. Kappa coefficients from ANN models with a single hidden layer, demonstrating a range of values from 0.05 to 0.77 (**Table 2.2**). However, more than 80% of ANN models presented Kappa coefficient values higher than 0.70 (**Figure 2.3**). The histogram illustrating the percentage of Kappa coefficient values from the ANN models demonstrated a mode value of 0.73, which occurred 86 times within the 378 models. The best Kappa coefficient was 0.77, relatively close to the mode value, considering the training rate of 0.2, training momentum of 0.4, and training threshold contribution of 0.7.

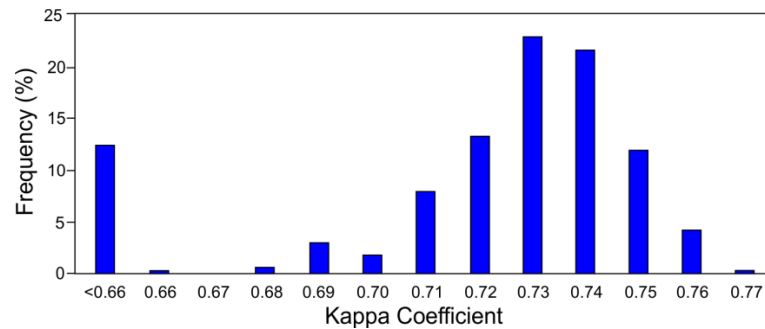


Figure 2.3. Histogram of percentage of kappa coefficients considering the different ANN models described in Table 2.

In the evaluation of internal parameters on the ANN performance, the average behavior under a wide range of empirical tests showed a trend that served as a guide for setting parameters (**Figure 2.4**). The training threshold contribution (**Figure 2.4a**) did not show significant contribution to the change in the Kappa coefficient values. The graph of the mean values of this parameter in the complete range from 0.1 to 0.9 evidenced a straight-line pattern (with the value of 0.67), demonstrating no influence in the results for the training set adopted in the research.

In contrast, the graphs of the mean values for the training momentum and training rate indicated a tendency to decrease the Kappa coefficient with increasing values for both parameters (**Figure 2.4b, c**). The training momentum showed a constant behavior for the Kappa coefficient between 0.05 and 0.6 followed by an abrupt drop in 0.8 and 0.9 (**Figure 2.4b**). Therefore, high momentum value may cause a divergent behavior with continued training, while lower values present better results. Similarly, the training rate exhibits initial rectilinear behavior between 0.005 and 0.2 followed by a decrease in the network performance from 0.2 to 0.8 (**Figure 2.4c**). The concomitant increase of learning rates and momentum values led to large steps within the weight space, resulting in a generalization of certain classes and a worsening of Kappa coefficients.

The four best-fit models adopted a constant training rate of 0.2 and a training moment predominantly of 0.4, showing results consistent with the behavior described by the average graphs (**Figure 2.5**). In contrast, training threshold contribution presented different values (0.5, 0.6, 0.7, 0.9) and did not influence the Kappa coefficient. **Table 2.3** lists the confusion matrices for the four best-fit models. The highest errors occurred between the double cropping systems, mainly Soybean/Cotton that were incorrectly marked as Soybean/Corn. The pasture class showed the highest amount of false positive, where different categories (annual single cultivation system, soybean/cotton, soybean/corn, central pivot) were considered as pasture. In contrast, the center pivot (triple cropping system) had a significant presence of false negatives. Commission and omission errors, highlighting the behaviors of the classes described by the confusion matrix (**Table 2.4**). The highest errors of commission were in the soybean / corn and pasture classes, with values higher than 30%. Meanwhile, the major errors of omission were in the central pivots with triple culture system, acquiring values superior to 30%. The areas of forest formation have the highest accuracy in the mapping.

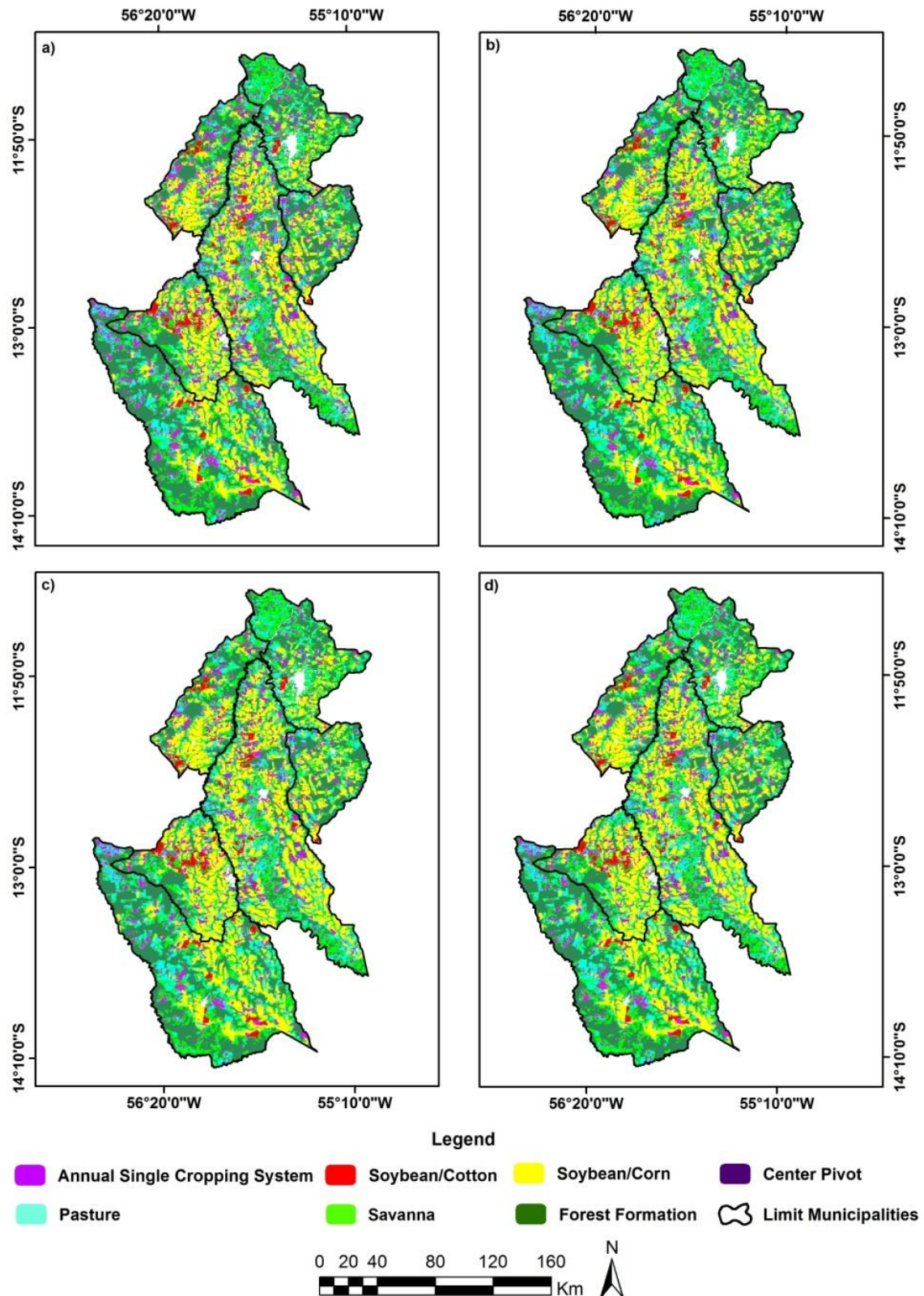


Figure 2.5 The four best-fit models using back-propagation neural networks considering value variations for training threshold contribution (v), training rate (η), and training momentum (μ): (a) first model ($v=0.7$, $\eta=0.2$, $\mu=0.4$); (b) second model ($v=0.6$, $\eta=0.2$, $\mu=0.4$); (c) third model ($v=0.5$, $\eta=0.2$, $\mu=0.4$); and (d) fourth model ($v=0.9$, $\eta=0.2$, $\mu=0.05$).

Table 2. 4. Commission and omission errors for the four best ANN classifications and classes (ASCS – Annual Single Cropping System, S/Ct – Soybean/Cotton, S/Cr – Soybean/Corn, CP– Center Pivot, P – Pasture, S – Savanna, FF – Forest Formation).

Classes	ANN - $v=0.7$, $\eta=0.2$, $\mu=0.4$		ANN - $v=0.6$, $\eta=0.2$, $\mu=0.4$		ANN - $v=0.5$, $\eta=0.2$, $\mu=0.4$		ANN - $v=0.9$, $\eta=0.2$, $\mu=0.05$	
	C %	O %	C %	O %	C %	O %	C %	O %
ASCS	17.70	7.00	17.11	6.33	18.18	7.00	17.26	7.33
S/Ct	7.63	27.33	7.63	27.33	8.02	27.33	6.17	29.00
S/Cr	33.23	27.67	33.33	28.00	33.64	28.33	34.76	28.67
CP	11.90	38.33	13.27	39.00	13.27	39.00	13.94	40.33
P	33.74	28.00	33.23	27.00	33.13	27.33	34.12	26.00
S	20.23	9.33	21.05	10.00	21.18	10.67	20.58	8.67
FF	7.43	0.33	7.50	1.33	8.10	1.67	7.21	1.33

Table 2.5 demonstrates the McNemar test results among the pairs of the four best ANN classifications considering the different combinations of parameters (v , η , and μ). At a significance level of 5% ($p = 0.05$), most of the comparisons presented calculated χ^2 values above the tabulated χ^2 value (3.841), rejecting the null hypothesis and confirming that the classifications are distinct. Therefore, the best ANN result ($v=0.7$, $\eta=0.2$, $\mu=0.4$) has a superiority in relation to the other methods, despite having values close to Kappa indices. Two pairs between ANN $\square v=0.6$, $\eta=0.2$, $\mu=0.4$) with ANN ($v=0.5$, $\eta=0.2$, $\mu=0.4$) and ANN ($v=0.9$, $\eta=0.2$, $\mu=0.05$) had equivalent results.

Table 2. 5. McNemar's test for the four best ANN classifications considering changes to the parameters: training threshold contribution (v), training rate (η), and training momentum (μ). The values marked in red have calculated χ^2 values higher than the critical value at a significance level of 5% (3.841).

Classification 1 x Classification 2	f_{11}	f_{12}	f_{21}	f_{22}	total	χ^2 result
($v=0.7$, $\eta=0.2$, $\mu=0.4$) X ($v=0.6$, $\eta=0.2$, $\mu=0.4$)	1560	18	50	472	2100	15.06
($v=0.7$, $\eta=0.2$, $\mu=0.4$) X ($v=0.5$, $\eta=0.2$, $\mu=0.4$)	1555	23	49	473	2100	9.39
($v=0.7$, $\eta=0.2$, $\mu=0.4$) X ($v=0.9$, $\eta=0.2$, $\mu=0.05$)	1555	23	64	458	2100	19.32
($v=0.6$, $\eta=0.2$, $\mu=0.4$) X ($v=0.5$, $\eta=0.2$, $\mu=0.4$)	1601	9	3	487	2100	3.00
($v=0.6$, $\eta=0.2$, $\mu=0.4$) X ($v=0.9$, $\eta=0.2$, $\mu=0.05$)	1588	22	31	459	2100	1.53
($v=0.5$, $\eta=0.2$, $\mu=0.4$) X ($v=0.9$, $\eta=0.2$, $\mu=0.05$)	1583	21	36	460	2100	3.95

2.3.3. Results of time-series classifications using Support Vector Machine

The spatial distributions of SVM results considering the different kernel functions (linear, polynomial, RBF, and sigmoid), which present very similar patterns. SVM classifications yielded worse results than the best ANN models (**Figure 2.6**). The SVM model with best accuracy indexes was the linear kernel function with the Overall accuracy of 78.90% and Kappa coefficient of 0.75, while the worse result was the

polynomial degree of 3 with the Overall accuracy of 77.57% and Kappa coefficient of 0.73. Therefore, the results of the accuracy indices are very close and should be evaluated by the McNemar's test to confirm the distinction between the methods.

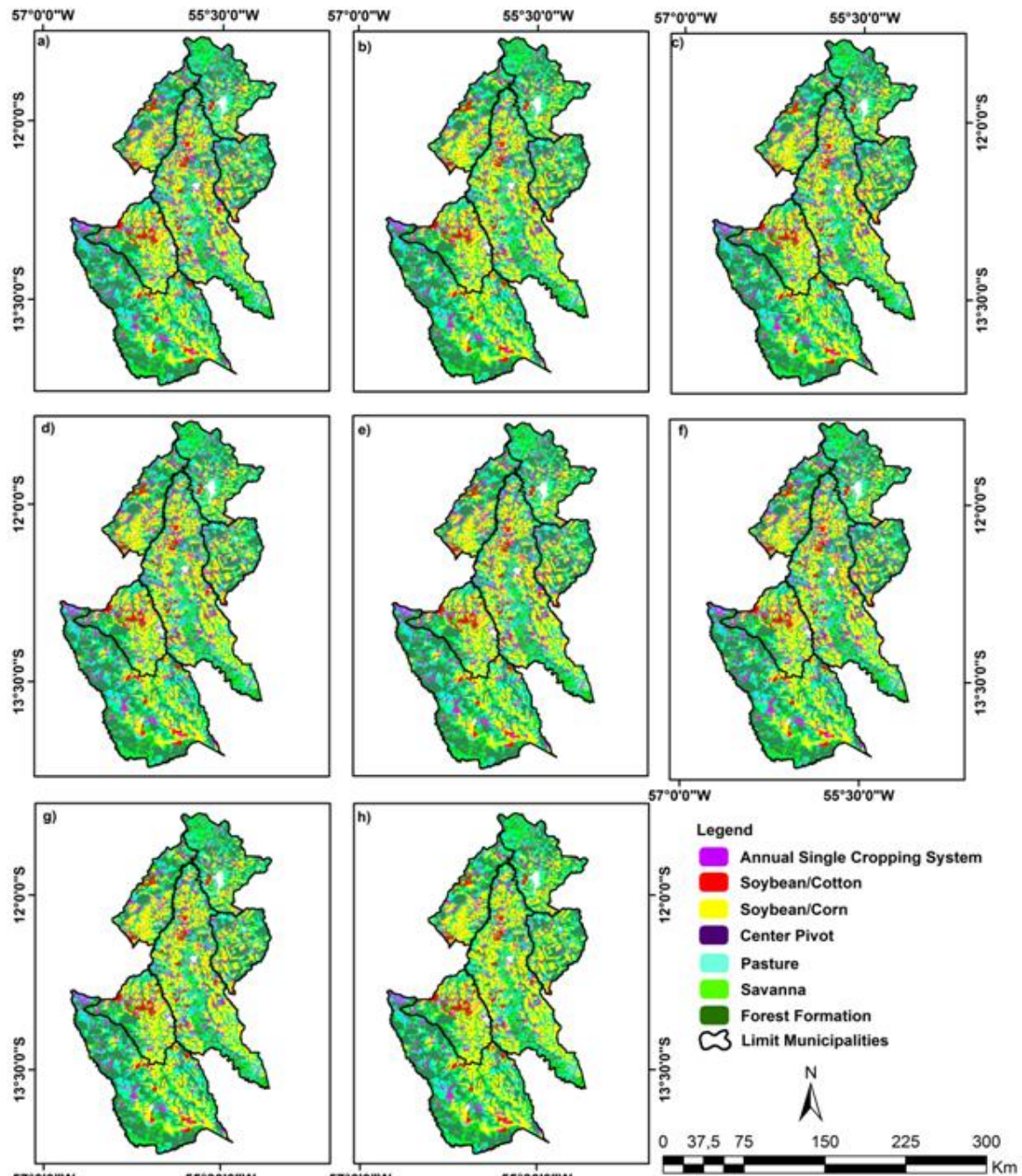


Figure 2.6. Classified images using Support Vector Machine using the following kernel types: (a) linear; (b) polynomial of degree 2; (c) polynomial of degree 3; (d) polynomial of degree 4; (e) polynomial of degree 5; (f) polynomial of degree 6; (g) RBF; and (h) sigmoid.

Table 2.6 lists the confusion matrices for the eight SVM models. Like ANN models, the highest errors that occurred were Soybean/Cotton as Soybean/Corn, annual single cropping systems as pasture, and Soybean/Corn as pasture. Therefore, the commission errors were significant in all methods for soybean/corn and pasture,

The different kernel functions showed omission-error values above 30% for center-pivot class, where these were misclassified as Soybean/Cotton, Soybean/Corn, and pasture. There were no systematic errors with natural vegetation types (forest formation and savanna). The results between the two models SVM and ANN possessed similar patterns, but the best ANN models provided a decrease in the intensity errors. The crop mapping showed that the double cropping systems of soybean/corn predominated in the six municipalities analyzed, while the soybean/cotton was cultivated in a smaller area. Single and triple cropping system was not significant in the study area (**Figures 2.5 and 2.6**).

Table 2. 7. Commission and omission errors for the different SVM methods and classes (ASCS – Annual Single Cropping System, S/Ct – Soybean/Cotton, S/Cr – Soybean/Corn, CP– Center Pivot, P – Pasture, S – Savanna, FF – Forest Formation).

Classes	Linear SVM		Polynomial SVM		RBF SVM		Sigmoid SVM	
	C %	O %	C %	O %	C %	O %	C %	O %
ASCS	19.70	11.67	16.41	27.0	15.95	28.0	14.17	27.33
S/Ct	16.08	28.67	14.45	27.0	15.06	26.67	14.67	26.33
S/Cr	38.13	38.33	35.65	32.0	35.11	31.0	36.11	31.0
CP	9.69	41.0	9.31	38.33	7.46	38.0	3.5	35.67
P	35.45	18.67	42.21	19.67	42.21	19.67	43.51	21.67
S	14.47	9.33	15.36	10.0	15.53	9.33	18.21	11.67
FF	7.41	0.0	7.69	0.0	7.69	0.0	8.05	1.0

Table 2.8 demonstrates the McNemar's test among the SVM classifications using the different kernel functions. The results showed that the classifications presented equivalent results at a significance level of 5%, not showing a superiority of one method over the other. The only exception was the sigmoid function, in which the test confirmed its inferiority in relation to the functions: linear, polynomial degree of 4 and 6. Therefore, the SVM results were uniform with the different kernels functions (except sigmoid).

Finally, the McNemar's test between the best ANN model ($v=0.7$, $\eta=0.2$, $\mu=0.4$) and the SVM models demonstrate an equivalence of the results, where the calculated χ^2 values are lower than the critical χ^2 with a significance of 5% (**Table 2.9**). Therefore, the equality between the best results of the two methods evidences that the two procedures can be adopted without loss of quality of the soybean, corn and cotton mapping in the study region.

Table 2. 8. McNemar's test among the different SVM models. The values marked in red have calculated χ^2 values higher than the critical value at a significance level of 5% (3.841).

Classification 1 x Classification 2	f ₁₁	f ₁₂	f ₂₁	f ₂₂	total	χ^2 result
Linear X Polynomial degree of 2	1517	68	51	464	2100	2.428571
Linear X Polynomial degree of 3	1526	59	49	466	2100	0.925926
Linear X Polynomial degree of 4	1539	46	43	472	2100	0.101124
Linear X Polynomial degree of 5	1539	46	38	477	2100	0.761905
Linear X Polynomial degree of 6	1552	33	34	481	2100	0.014925
Linear X RBF	1514	71	60	455	2100	0.923664
Linear X Sigmoid	1484	101	74	441	2100	4.165714
Polynomial degree of 2 X 3	1552	16	23	509	2100	1.25641
Polynomial degree of 2 X 4	1545	23	37	495	2100	3.266667
Polynomial degree of 2 X 5	1534	34	43	489	2100	1.051948
Polynomial degree of 2 X 6	1531	37	55	477	2100	3.521739

Polynomial degree of 2 X RBF	1561	7	13	519	2100	1.8000
Polynomial degree of 2 X Sigmoid	1519	49	39	493	2100	1.136364
Polynomial degree of 3 X 4	1559	16	23	502	2100	1.25641
Polynomial degree of 3 X 5	1548	27	29	496	2100	0.071429
Polynomial degree of 3 X 6	1543	32	43	482	2100	1.613333
Polynomial degree of 3 X RBF	1549	26	25	500	2100	0.019608
Polynomial degree of 3 X Sigmoid	1511	64	47	478	2100	2.603604
Polynomial degree of 4 X 5	1566	16	11	507	2100	0.925926
Polynomial degree of 4 X 6	1556	26	30	488	2100	0.285714
Polynomial degree of 4 X RBF	1544	38	30	488	2100	0.941177
Polynomial degree of 4 X Sigmoid	1509	73	49	469	2100	4.721312
Polynomial degree of 5 X 6	1563	14	23	500	2100	2.189189
Polynomial degree of 5 X RBF	1533	44	41	482	2100	0.105882
Polynomial degree of 5 X Sigmoid	1498	79	60	463	2100	2.597122
Polynomial degree of 6 X RBF	1529	57	45	469	2100	1.411765
Polynomial degree of 6 X Sigmoid	1497	89	61	453	2100	5.226666
RBF X Sigmoid	1530	44	28	498	2100	3.555556

Table 2. 9. McNemar's test between the best ANN model ($v=0.7$, $\eta=0.2$, $\mu=0.4$) and SVM models.

Classification 1 x Classification 2	f ₁₁	f ₁₂	f ₂₁	f ₂₂	total	χ^2 result
ANN ($v=0.7$, $\eta=0.2$, $\mu=0.4$) X SVM (Linear)	1506	72	79	443	2100	0.324503
ANN ($v=0.7$, $\eta=0.2$, $\mu=0.4$) X SVM (Polynomial degree of 2)	1491	87	77	445	2100	0.609756
ANN ($v=0.7$, $\eta=0.2$, $\mu=0.4$) X SVM (Polynomial degree of 3)	1501	77	74	448	2100	0.059603
ANN ($v=0.7$, $\eta=0.2$, $\mu=0.4$) X SVM (Polynomial degree of 4)	1504	74	78	444	2100	0.105263
ANN ($v=0.7$, $\eta=0.2$, $\mu=0.4$) X SVM (Polynomial degree of 5)	1504	74	73	449	2100	0.006803
ANN ($v=0.7$, $\eta=0.2$, $\mu=0.4$) X SVM (Polynomial degree of 6)	1518	60	68	454	2100	0.5
ANN ($v=0.7$, $\eta=0.2$, $\mu=0.4$) X SVM (RBF)	1495	83	79	443	2100	0.098765

2.4. DISCUSSION

In Central Brazil, the multiple cropping agriculture using farm mechanization and high amount of chemical fertilizer has been a key factor in economic development. Although the cropping intensity varies annually due to natural or economic factors, the growth in the study area has been steady in recent years. Traditionally, the procedures to estimate the agricultural area and cropping intensity are based on fieldwork and agricultural census, which is a very slow and costly process. Alternatively, studies draw attention to the processing of high-temporal resolution datasets that provides the mapping of crop phenology and cropping intensity, a critical factor for production, in a less costly and faster way (SAKAMOTO et al., 2005, 2010; WARDLOW; EGBERT; KASTENS, 2007).

In this research, we tested the back-propagation neural network and SVM methods for classification of S-G-smoothed NDVI time series. The choice of the internal parameters and geometry of back-propagation neural networks is an important factor for process optimization and mapping accuracy (MAIER; DANDY, 1998). These values are estimated from an investigation procedure to obtain information about the network operation, serving as a guide for future classifications in the study region (HEERMANN; KHAZENIE, 1992). The 378 back-propagation models with a single hidden layer were tested to determine the parameters that provided high accuracy. The main evaluated parameters were training rate, training momentum, and training threshold contribution. The back-propagation algorithm decreases efficiency with too large steps for training rate (>0.2) and training momentum (>0.6). The different values of training threshold contribution do not present significant variation in the Kappa coefficients. Consequently, analytical techniques to determine the inputs to ANN multivariate models bring advantages in mapping accuracy. The McNemar's test demonstrated a statistically significant effect in the thematic maps with highest Kappa index from ANN models.

Except for the sigmoid function, all other SVM models using different kernel functions (linear, polynomial and RBF) demonstrated by the McNemar's test to accept the null hypothesis of no processing effect. Furthermore, these SVM models have equivalence with the best ANN models, evidencing that both classifiers in the study area allow the evaluation of spatiotemporal patterns of agricultural and natural areas with reasonable estimation accuracy. Shao & Lunetta (SHAO; LUNETTA, 2012) describe a better performance of SVM method in relation to the neural network algorithm for the land-cover classification using MODIS time-series data in North Carolina and Virginia (USA). Probably, the similar behavior between the two methods found in the present research may be due to the high number of tests performed with the input parameters of the ANN method. Therefore, the use of SVM is advisable due to its equivalence of results using different kernel functions (except the sigmoid), which equates to the best ANN model considering a broad set of tests.

Past studies have been conducted in this region for the mapping of crops using MODIS images, considering different land-use classes and methodologies that make it difficult to compare directly with the present study (GALFORD et al., 2008; ARVOR et al., 2011; BERNARDES et al., 2011; GUSSO et al., 2014). The researchers developed by Galford et al. (2008), Bernardes et al. (2011), and Gusso et al. (2014) had a lower range of classes. Galford et al. (2008) aimed to separate only single and double cropping patterns from the number of local maximums during one year in the wavelet-smoothed EVI time series. Bernardes et al., (2011) developed a methodology based on principal component analysis and ISOSEG unsupervised classification for only soybean mapping in Mato Grosso State, using MODIS daily and 16-day composite products. Similarly, Gusso et al., (2014) mapped the soybean and non-soybean classes to the Mato Grosso state utilizing a decision tree system based on different data: MODIS time series (Minimum and Maximum mean EVI image), precipitation, terrain slope from the Shuttle Radar Topography Mission data, and soil classes.

In the mapping of natural vegetation and agricultural crops in the region of Mato Grosso state, Arvor et al., (2011) used a two-step classification of the MODIS time series, where the first stage separated the natural areas, pastures and agricultural areas; and the second stage individualized the types of crops within the agricultural areas. Both classification steps were based on feature extraction, maximum likelihood, and segmentation. The authors in the first classification obtained a Kappa index of 0.80, while

in the second classification they obtained a Kappa index of 0.67. The accuracy values obtained separately for the two steps make it difficult to compare with the present research, but the error patterns are similar with high values between pasture and savanna.

The class of triple cropping systems was not addressed in the surveys using time-series imagery in the study region. Although the triple cropping systems had distinct temporal signatures, we verified high omission-error values for the different methods. Biradar and Xiao (2011) also had trouble mapping the triple cropping in India, considering the temporal signature analysis of three vegetation indices (NDVI, EVI, and LSWI) from MODIS data at the 500-m spatial resolution and 8-day temporal resolution. These authors obtained 81% agreement in double cropping and 36% agreement in triple cropping, wherein this low accuracy was due to the small areas and short vegetable growing cycle.

2.5. CONCLUSIONS

The present research evaluated the performance of back-propagation neural network and Support Vector Machine to map multiple cropping areas using NDVI-MODIS time series data at 250-m spatial resolution. Time-series classification is efficient for agricultural mapping because it allows describing the phenological variation of the crop that is hard to obtain purely considering the spectral signatures. In the crop phenology detection, various models were evaluated using the back-propagation neural network with different input parameters and Support Vector Machine with different kernel function. Thus, we adopted a test series to find the optimal settings for the evaluated methods. The back-propagation algorithm becomes more efficient with lower values of training rate (<0.2) and training momentum (<0.6). Training threshold contribution had little influence on the result. Fine-tuning of back-propagation parameter set provided an optimal result with Kappa coefficient of 0.77 and Overall Accuracy of 80.28%. The McNemar test considering ANN models with better Kappa indices rejected the null hypothesis of no effect of the parameters in these classifications. However, the McNemar test among the SVM models using the different kernel function (linear, polynomial and RBF) and the best ANN result showed that, regardless of the classification method, the results are statistically equivalent. Given the results, time-series classifications using Neural Network and SVM are appropriate to identify the crops in Mato Grosso State, even considering the limited spatial resolution of the MODIS sensor data. The Brazilian government could use this procedure based on multitemporal remotely sensed data (CONAB and IBGE) to define production area giving information to manage pricing strategies replacing partially the time consuming and high-cost farmers interviews.

Acknowledgments: The National Counsel of Technological and Scientific Development (CNPq) funded this study. The authors are grateful for financial support from CNPq fellowship (Osmar Abílio de Carvalho Júnior and Renato Fontes Guimarães). Special thanks are given to Soybean and Corn Producers Association (APROSOJA), Brazilian Agricultural Research Corporation (EMBRAPA) Agrosilvopastoral and National Supply Company (CONAB) for providing data and fieldwork support. The authors thank the researchers Sandro Nunes de Oliveira and Nickolas Castro Santana at the University of Brasília and the agronomist Gabriel Ponce Alves at the Company Soil Plant for collaboration in the field works. Finally, the authors acknowledge the contribution from anonymous reviewers.

REFERENCES

- ABADE, N.; JÚNIOR, O.; GUIMARÃES, R.; DE OLIVEIRA, S. Comparative Analysis of MODIS Time-Series Classification Using Support Vector Machines and Methods Based upon Distance and Similarity Measures in the Brazilian Cerrado-Caatinga Boundary. *Remote Sensing*, v. 7, n. 9, p. 12160–12191, 18 set. 2015.
- AGUIAR, D. A. De; FRIEDRICH, B.; RUDORFF, T.; RIZZI, R.; SHIMABUKURO, Y. E. Monitoramento da colheita da cana-de-açúcar por meio de imagens MODIS. *Revista Brasileira de Cartografia*, v. 60, p. 375–383, 2008.
- ALCANTARA, C.; KUEMMERLE, T.; PRISHCHEPOV, A. V.; RADELOFF, V. C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. **Remote Sensing of Environment**, v. 124, p. 334–347, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.rse.2012.05.019>>.
- ANTUNES, J. F. G.; LAMPARELLI, R. A.; RODRIGUES, L. H. A. Avaliação da dinâmica do cultivo da cana-de-açúcar no estado de são paulo por meio de perfis temporais de dados modis. *Revista Engenharia Agrícola*, v. 4430, p. 1127–1136, 2015.
- ARVOR, D.; JONATHAN, M.; MEIRELLES, M. S. P.; DUBREUIL, V.; DURIEUX, L. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *International Journal of Remote Sensing*, v. 32, n. 22, p. 7847–7871, 2011.
- ATKINSON, P. M.; JEGANATHAN, C.; DASH, J.; ATZBERGER, C. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment*, v. 123, p. 400–417, 2012.
- ATZBERGER, C. Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sensing*, v. 5, n. 2, p. 949–981, 2013.
- ATZBERGER, C.; REMBOLD, F. Mapping the spatial distribution of winter crops at sub-pixel level using AVHRR NDVI time series and neural nets. *Remote Sensing*, v. 5, n. 3, p. 1335–1354, 2013.
- BAHIA, R. B. C.; MARTINS-NETO, M. A.; BARBOSA, M. S. C.; PEDREIRA, A. J. Revisão Estratigráfica Da Bacia Dos Parecis – Amazônia. *Revista Brasileira de Geociências*, v. 36, n. 4, p. 692–703, 2006.
- BAHIA, R. B. C.; NETO, M. A. M. Análise da evolução tectonossedimentar da Bacia dos Parecis através de métodos potenciais. *Revista Brasileira de Geociências*, v. 37, n. 4, p. 639–649, 2007.
- BARROS, A.; DA SILVA, R.; CARDOSO, O.; FREIRE, F.; DE SOUZA JÚNIOR, J.; RIVETTI, M.; DA LUZ, D.; PALMEIRA, R.; TASSINARI, C. Projeto RADAMBRASIL, Levantamento de Recursos Naturais - Folha Cuiabá, SD. 21. Rio de Janeiro: Ministério das Minas e Energia, 1982.
- BECKER-RESHEF, I.; JUSTICE, C.; SULLIVAN, M.; VERMOTE, E.; TUCKER, C.; ANYAMBA, A.; SMALL, J.; PAK, E.; MASUOKA, E.; SCHMALTZ, J.; HANSEN, M.; PITTMAN, K.; BIRKETT, C.; WILLIAMS, D.; REYNOLDS, C.; DOORN, B.

Monitoring Global Croplands with Coarse Resolution Earth Observations: The Global Agriculture Monitoring (GLAM) Project. *Remote Sensing*, v. 2, n. 6, p. 1589–1609, 2010.

BERNARDES, T.; ADAMI, M.; FORMAGGIO, A. R.; MOREIRA, M. A.; FRANÇA, D. de A.; DE NOVAES, M. R. Imagens mono e multitemporais Modis para estimativa da área com soja no Estado de Mato Grosso. *Pesquisa Agropecuaria Brasileira*, v. 46, n. 11, p. 1530–1537, 2011.

BIRADAR, C. M.; XIAO, X. Quantifying the area and spatial distribution of double- and triple-cropping croplands in India with multi-temporal MODIS imagery in 2005. *International Journal of Remote Sensing*, v. 32, n. 2, p. 367–386, 2011.

CARVALHO JÚNIOR, O. A. de; DA SILVA, N. C.; DE CARVALHO, A. P. F.; COUTO JÚNIOR, A. F.; SILVA, C. R.; SHIMABUKURO, Y. E.; GUIMARÃES, R. F.; GOMES, R. A. T. Combining Noise-Adjusted Principal Components transform and median filter techniques for denoising MODIS temporal signatures. *Revista Brasileira de Geofísica*, v. 30, n. 2, p. 147–157, 2012.

CARVALHO JÚNIOR, O. A. de; GUIMARÃES, R. F.; SILVA, C.; GOMES, R. A. T. Standardized Time-Series and Interannual Phenological Deviation: New Techniques for Burned-Area Detection Using Long-Term MODIS-NBR Dataset. *Remote Sensing*, v. 7, n. 6, p. 6950–6985, 2015.

CASTRO, C. N. de. A agropecuária na região Centro-Oeste: limitações ao desenvolvimento e desafios futuros. *Texto para Discussão IPEA*, v. 1923, p. 1–4, 2014.

CASTRO JÚNIOR, P. R.; PAULO, U. D. S. Dinâmica da Água em Campos de Murundus do Planalto dos Parecis. p. 193, 2002.

CHEN, J.; JÖNSSON, P.; TAMURA, M.; GU, Z.; MATSUSHITA, B.; EKLUNDH, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment*, v. 91, n. 3–4, p. 332–344, 2004.

CONGALTON, R. G.; ODERWALD, R. G.; MEAD, R. A. Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques. *Photogrammetric Engineering and Remote Sensing*, v. 49, n. No. 12, p. 1671–1678, 1983.

COUTO JÚNIOR, A. F.; CARVALHO JÚNIOR, O. A.; MARTINS, É. D. S. Séries temporais MODIS aplicadas em sucessão de culturas de soja (*Glycine max* (L.) Merrill) e milho (*Zeamays* L.) em sistema de plantio direto. *Revista Brasileira de Cartografia*, v. 64, n. 3, p. 405–418, 2012.

COUTO JUNIOR, A. F.; CARVALHO JUNIOR, O. A.; MARTINS, E. S.; SANTANA, O. A.; SOUZA, V. V.; ENCONAS, J. I. Tratamento De Ruídos E Caracterização De Fisionomias Do Cerrado Utilizando Séries Temporais Do Sensor Modis. *Revista Árvore*, v. 35, n. 3, p. 699–705, 2011.

COUTO JÚNIOR, A. F.; DE CARVALHO JUNIOR, O. A.; MARTINS, E. S.; GUERRA, A. F. Phenological characterization of coffee crop (*Coffea Arabica* l.) from MODIS time series. *Revista Brasileira de Geofísica*, v. 31, n. 4, p. 569–578, 2013.

CUTRIM, A. O.; RUIZ, A. S.; LIPORONI, L. M.; MEDEIROS, F. a. De; BARROSO, U. C.; NASCIMENTO, A. L. Sondagem elétrica vertical aplicada em pesquisa hidrogeológica na bacia do Parecis, MT. *Revista Brasileira de Geofísica*, v. 25, n. 2, p. 131–140, jun. 2007.

DWYER, J.; SCHMIDT, G. The MODIS Reprojection Tool. In: *Earth Science Satellite Remote Sensing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. p. 162–177.

FOODY, G. M. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, v. 80, n. 1, p. 185–201, 2002.

FOODY, G. M. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering & Remote Sensing*, v. 70, n. 5, p. 627–633, 2004.

FRANCH, B.; VERMOTE, E. F.; BECKER-RESHEF, I.; CLAVERIE, M.; HUANG, J.; ZHANG, J.; JUSTICE, C.; SOBRINO, J. A. Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information. *Remote Sensing of Environment*, v. 161, p. 131–148, 2015.

GALFORD, G. L.; MUSTARD, J. F.; MELILLO, J.; GENDRIN, A.; CERRI, C. C.; CERRI, C. E. P. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment*, v. 112, n. 2, p. 576–587, 2008.

GENG, L.; MA, M.; WANG, X.; YU, W.; JIA, S.; WANG, H. Comparison of eight techniques for reconstructing multi-satellite sensor time-series NDVI data sets in the heihe river basin, China. *Remote Sensing*, v. 6, n. 3, p. 2024–2049, 2014.

GUMMA, M. K.; MOHANTY, S.; NELSON, A.; ARNEL, R.; MOHAMMED, I. A.; DAS, S. R. Remote sensing based change analysis of rice environments in Odisha, India. *Journal of Environmental Management*, v. 148, p. 31–41, 2015.

GUMMA, M. K.; NELSON, A.; THENKABAIL, P. S.; SINGH, A. N. Mapping rice areas of South Asia using MODIS multitemporal data. *Journal of Applied Remote Sensing*, v. 5, n. 1, p. 53547, 2011.

GUSSO, A.; ARVOR, D.; RICARDO DUCATI, J.; VERONEZ, M. R.; DA SILVEIRA, L. G. Assessing the modis crop detection algorithm for soybean crop area mapping and expansion in the Mato Grosso state, Brazil. *The Scientific World Journal*, v. 2014, n. 1, 2014.

HARTFIELD, K. A.; LANDAU, K. I.; VAN LEEUWEN, W. J. D. Fusion of high resolution aerial multispectral and lidar data: Land cover in the context of urban mosquito habitat. *Remote Sensing*, v. 3, n. 11, p. 2364–2383, 2011.

HEERMANN, P. D.; KHAZENIE, N. Classification of multispectral remote sensing data using a back-propagation neural network. *IEEE Transactions on Geoscience and Remote Sensing*, v. 30, n. 1, p. 81–88, 1992.

HIRD, J. N.; MCDERMID, G. J. Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment*, v. 113, n. 1, p. 248–258, 2009.

HÜTTICH, C.; GESSNER, U.; HEROLD, M.; STROHBACH, B. J.; SCHMIDT, M.; KEIL, M.; DECH, S. On the suitability of MODIS time series metrics to map vegetation types in dry savanna ecosystems: A case study in the Kalahari of NE Namibia. *Remote Sensing*, v. 1, n. 4, p. 620–643, 2009.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Levantamento Sistemático da Produção Agrícola - LSPA. 2015. Disponível em: <<http://www.ibge.gov.br/>>. Acesso em: 1 maio. 2017.

JÖNSSON, P.; EKLUNDH, L. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 40, n. 8, p. 1824–1832, 2002.

KASTENS, J. H.; KASTENS, T. L.; KASTENS, D. L. A.; PRICE, K. P.; MARTINKO, E. A.; LEE, R. Y. Image masking for crop yield forecasting using AVHRR NDVI time series imagery. *Remote Sensing of Environment*, v. 99, n. 3, p. 341–356, 2005.

LUNETTA, R. S.; KNIGHT, J. F.; EDIRIWICKREMA, J.; LYON, J. G.; WORTHY, L. D. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment*, v. 105, n. 2, p. 142–154, 2006.

LUNETTA, R. S.; SHAO, Y.; EDIRIWICKREMA, J.; LYON, J. G. Monitoring agricultural cropping patterns across the Laurentian Great Lakes Basin using MODIS-NDVI data. *International Journal of Applied Earth Observation and Geoinformation*, v. 12, n. 2, p. 81–88, 2010.

MAIER, H. R.; DANDY, G. C. The effect of internal parameters and geometry on the performance of back-propagation neural networks: An empirical study. *Environmental Modelling and Software*, v. 13, n. 2, p. 193–209, 1998.

MAS, J. F.; FLORES, J. J. The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, v. 29, n. 3, p. 617–663, 2008.

MCNEMAR, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, v. 12, n. 2, p. 153–157, 1947.

MOUNTRAKIS, G.; IM, J.; OGOLE, C. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and*, 2011.

OZDOGAN, M. The spatial distribution of crop types from MODIS data: Temporal unmixing using Independent Component Analysis. *Remote Sensing of Environment*, v. 114, n. 6, p. 1190–1204, 2010.

PAN, Y.; LI, L.; ZHANG, J.; LIANG, S.; ZHU, X.; SULLA-MENASHE, D. Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. *Remote Sensing of Environment*, v. 119, p. 232–242, 2012.

PRINGLE, M. J.; DENHAM, R. J.; DEVADAS, R. Identification of cropping activity in central and southern Queensland, Australia, with the aid of MODIS MOD13Q1 imagery. *International Journal of Applied Earth Observation and Geoinformation*, v. 19, n. 1, p. 276–285, 2012.

REN, J.; CHEN, Z.; ZHOU, Q.; TANG, H. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. *International Journal of Applied Earth Observation and Geoinformation*, v. 10, n. 4, p. 403–413, 2008.

REZENDE, G. C. de. Ocupação Agrícola e Estrutura Agrária no Cerrado: O Papel do Preço da Terra, dos Recursos Naturais e da Tecnologia. Texto para Discussão IPEA, v. 913, p. 1–28, 2002.

RICHARDS, J. A. Supervised Classification Techniques. *Remote Sensing Digital Image Analysis: An Introduction*, p. 494, 2013.

ROSA, V. G. C. da; MOREIRA, M. A.; RUDORFF, B. F. T.; ADAMI, M. Estimativa da produtividade de café com base em um modelo agrometeorológico-espectral. *Pesquisa Agropecuária Brasileira*, v. 45, n. 12, p. 1478–1488, 2010.

ROUSE, J. W.; HASS, R. H.; SCHELL, J.; DEERING, D. Monitoring vegetation systems in the great plains with ERTS Third Earth Resources Technology Satellite (ERTS) symposium, 1973. .

SAKAMOTO, T.; GITELSON, A. A.; ARKEBAUER, T. J. MODIS-based corn grain yield estimation model incorporating crop phenology information. *Remote Sensing of Environment*, v. 131, p. 215–231, 2013.

SAKAMOTO, T.; WARDLOW, B. D.; GITELSON, A. Detecting spatiotemporal changes of corn developmental stages in the US corn belt using Detecting Spatiotemporal Changes of Corn Developmental Stages in the U . S . Corn Belt Using MODIS WDRVI Data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 49, n. 6, p. 1926–1936, 2011.

SAKAMOTO, T.; WARDLOW, B. D.; GITELSON, A. A.; VERMA, S. B.; SUYKER, A. E.; ARKEBAUER, T. J. A Two-Step Filtering approach for detecting maize and soybean phenology with time-series MODIS data. *Remote Sensing of Environment*, v. 114, n. 10, p. 2146–2159, 2010.

SAKAMOTO, T.; YOKOZAWA, M.; TORITANI, H.; SHIBAYAMA, M.; ISHITSUKA, N.; OHNO, H. A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, v. 96, n. 3–4, p. 366–374, 2005.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, v. 36, n. 8, p. 1627–1639, 1964.

SCHÖLKOPF, B. Learning with kernels. *Journal of the Electrochemical Society*, v. 129, n. November, p. 2865, 2002.

SHAO, Y.; LUNETTA, R. S. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points.

ISPRS Journal of Photogrammetry and Remote Sensing, v. 70, n. June 2012, p. 78–87, 2012.

SHAO, Y.; LUNETTA, R. S.; WHEELER, B.; IIAMES, J. S.; CAMPBELL, J. B. An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data. *Remote Sensing of Environment*, v. 174, p. 258–265, 2016.

SHMILOVICI, A. Support Vector Machines. In: *Data Mining and Knowledge Discovery Handbook*. New York: Springer-Verlag, 2005. p. 257–276.

TELUGUNTLA, P.; RYU, D.; GEORGE, B.; WALKER, J.; MALANO, H. Mapping Flooded Rice Paddies Using Time Series of MODIS Imagery in the Krishna River Basin, India. *Remote Sensing*, v. 7, n. 7, p. 8858–8882, 2015.

WARDLOW, B. D.; EGBERT, S. L.; KASTENS, J. H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment*, v. 108, n. 3, p. 290–310, 2007.

WESKA, R. K. Uma síntese do Cretáceo superior Mato-Grossense. *Geociencias*, v. 25, n. 1, p. 71–81, 2006.

XAVIER, A. C.; RUDORFF, B. F. T.; SHIMABUKURO, Y. E.; MIURA, L.; BERKA, S. International Journal of Remote Multi - temporal analysis of MODIS data to classify sugarcane crop. *International Journal of Remote Sensing*, v. 27, n. 4, p. 755–768, 2006.

XIAO, X.; BOLES, S.; FROLKING, S.; LI, C.; BABU, J. Y.; SALAS, W.; MOORE, B. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment*, v. 100, n. 1, p. 95–113, 2006.

XIAO, X.; BOLES, S.; LIU, J.; ZHUANG, D.; FROLKING, S.; LI, C.; SALAS, W.; MOORE, B. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment*, v. 95, n. 4, p. 480–492, 2005.

ZHANG, X.; FRIEDL, M. A.; SCHAAF, C. B.; STRAHLER, A. H.; HODGES, J. C. F.; GAO, F.; REED, B. C.; HUETE, A. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, v. 84, n. 3, p. 471–475, 2003.

CAPÍTULO 3 – WHEAT PLANTED AREA DETECTION FROM THE MODIS NDVI TIME SERIES CLASSIFICATION USING THE NEAREST NEIGHBOR METHOD CALCULATED BY THE EUCLIDEAN DISTANCE AND COSINE SIMILARITY MEASURES

ABSTRACT

Wheat is one of the world's leading cereal crops, with the largest volume crop traded internationally. Accurate information on wheat production at a regional and global scale is useful for territorial management, economic planning, and food security. In this context, different remote sensing techniques using time series have been developed to estimate planted area, agricultural monitoring, and production forecast. In Brazil, wheat production is concentrated in the southern region of the country due to more favorable temperature and adapted cultivars. However, climatic oscillations cause significant spatiotemporal variations, resulting in changes in sowing time and plant growth stages. These varieties of the temporal signatures are a problem to be considered in the wheat detection by remote sensor data in this region. This research aims to detect the wheat crop in the Northwest region of Rio Grande do Sul (Brazil) using MODIS NDVI time series. The proposed methodology was composed of the following steps: (a) acquisition of MODIS images; (b) data smoothing using the Savitzky-Golay filter; (c) selection of reference temporal signatures; (d) classification based on nearest neighbor (a specific case of the K-NN method) from the similarity and distance metrics; and (e) and accuracy analysis. The classification based on nearest neighbor presented two steps: (a) generation of metric images (Euclidean distance and cosine similarity); and (b) definition of the best threshold value to individualize the wheat mask. Different temporal profiles of the wheat crop varied according to the agricultural calendars of each region. The nearest neighbor classification (K-NN) was operational and viable for the wheat crop mapping. The minimum distance method (Kappa index of 0.75) obtained a result equivalent to that of cosine similarity (Kappa index of 0.74) as attested by the McNemar test.

Keywords: crop classification, k-nearest neighbors, optimal threshold detection, phenology, Brazil.

3.1. INTRODUCTION

Wheat is the most cultivated cereal in the world, occupying an annual planted area of more than 220 million hectares in diverse climatic conditions and geographic regions (Shiferaw et al. 2013). Also, wheat has high relevance in the world's nutrition due to its quality and quantity of proteins and variety of derived products (Sun et al. 2012), being responsible for about 20% of the total dietary calories and proteins worldwide (Shiferaw et al. 2013). Atchison et al. (2010) examined 10,235 supermarket items in Australia and found that 20.7 percent contained wheat, with 1977 items in food and 144 items in non-food items, while considering only the food products the presence of wheat increased to about 29.5%. This cereal lead world trade in agricultural commodities and is the leading food import product of developing nations (Dixon et al. 2009). During the last century, wheat productivity increased from the improvement of several factors: wheat genetics, mechanization, pest and disease control, agricultural management and practice (Dixon et al. 2006). However, global demand for wheat tends to increase significantly due to the growth of urban populations that consume

more animal products and choose wheat-based products that are easy to prepare (Pingali 2007; Lobell and Burke 2010). Moreover, a future concern is that global warming may represent a fall in world wheat production by 6% for each degree Celsius of temperature rise (Asseng et al. 2014). Consequently, future food security depends on the continuity of productivity gains, addressing the threats of climate change, and establishing a monitoring system of production and demand.

Therefore, wheat combined with rice are internationally essential products for the global food security and poverty reduction, which results in an effort to your monitoring and crop forecasting. The provision of reliable information on agricultural production is crucial for the world's policy-makers and food security at the national and global level. The traditional acquisition of information on agricultural production is carried out through national census. However, the acquisition of agricultural census data from household surveys (in situ) is expensive and requires a great deal of fieldwork and time. Furthermore, this approach generates data that contain aggregated information about plantation types and their areas, not allowing a detailed description of the spatial and temporal distribution of crops. Hence, there is a consensus that the use of remote sensing techniques enables an improvement in the acquisition of agricultural data, establishing an accurate spatiotemporal distribution of the crops, faster data acquisition, and lower cost. These conclusions have been demonstrated by researchers in different countries, such as India (Biggs et al. 2006), Pakistan (Dempewolf et al. 2014), China (Frolking et al. 1999), South and Southeast Asia (Xiao et al. 2006), Brazil (Junges and Fontana 2009), among others. Thus, several surveys using remote sensing have been developed to support a wide range of users (government agencies, researchers, and farmers) (Lobell and Asner 2004). The estimation of wheat production has many applications, such as the formulation of supply policies and the grain industry.

Remote sensing techniques for the agricultural crop mapping focus on time series processing, which identifies the cultivation types by their phenological phases, impossible from data of only one date (Pan et al. 2015; Hao et al. 2016). In this context, high temporal resolution-satellite data bring new possibilities for solving problems of crop phenology detection. The MODIS (Moderate Resolution Imaging Spectroradiometer) sensor has been widely used in the temporal analysis because of its high temporal resolution (quasi-daily repetition) and moderate spatial resolution (250m) (Justice et al. 1998). MODIS images and time series processing techniques have been widely used for different purposes: area estimation and spatial distribution (e.g., Ozdogan 2010; Pan et al. 2012); monitoring (e.g., Wenbo et al. 2007; Feng et al. 2009a, b); wheat leaf area index retrieval (e.g., Yi et al. 2008; Dong et al. 2016); gross primary productivity (e.g., Yan et al. 2009); forecasting wheat yields (e.g., Ren et al. 2008; Becker-Reshef et al. 2010; Wit et al. 2012; Kouadio et al. 2012).

In Brazil, the main winter crop is wheat with an annual production of more than six million tons distributed in the South, Southeast, and Central regions. However, Brazil is one of the world's largest importers of wheat, importing about half of its consumption. Traditionally, the primary supplier of wheat to Brazil is Argentina. The main wheat-growing region is Southern Brazil, concentrating 92% of the national production, where the State of Paraná is the most abundant producer (51% of national production) followed by Rio Grande do Sul (37% of national production) (IBGE, 2016). The concentration of wheat cultivation in Southern Brazil is due to the cold temperature and the development of cultivars adapted for this region (Canziani and Guimarães 2009).

However, expansion of wheat cultivation in Southern Brazil is limited by variations in annual production due to adverse climatic fluctuations during the crop cycle. The climate has a strong influence on the development of the wheat crop, conditioning its yield, economic viability, and industrial quality characteristics. Climatic risk factors in the state of Rio Grande do Sul have significant implication in the wheat plantations. Cunha et al. (2001) generated

maps of the most suitable wheat-sowing time in the State of the Rio Grande do Sul from the following risk maps: frost during flowering and excess rainfall before harvest. The results show that the sowing time has a significant delay between the regions. In the northwest region of the State, sowing time begins earlier due to its warmer temperatures (May 1), while the northeastern part with higher altitudes and lower temperatures may occur with a lag time of three months (July 31). This variation is also identified in the state of Paraná, neighbor to the state of Rio Grande do Sul. Silva et al. (2011) identified the best sowing dates for wheat cultivars in two regions of the state of Paraná, where grain yield maximization in the Guarapuava region occurred with sowing in July, while in the Palotino region in April. Therefore, wheat crop phenology has a pronounced variation in South Brazil, since the climatic variations and plant development stages are not uniform in the spatial-temporal scale.

The present research aims to distinguish the wheat crop in the Northwest region of Rio Grande do Sul (Brazil) using MODIS time-series classification by the nearest neighbor method, a specific case of the K-NN method. Complementarily, the research seeks to test and compare the metrics of cosine similarity and Euclidean distance in the calculation of the nearest neighbor.

3.2. STUDY AREA

The study area is the Northwest region of Rio Grande do Sul, which includes 216 municipalities divided into 13 Microregions, where wheat production is concentrated within the state (Camponogara et al. 2015) (**Figure 3.1**). According to Köppen's climatic classification, the region is in the Temperate Climatic Zone (C), humid fundamental climate (f), subtropical specific variety (Cfa), with well-distributed precipitation during the year and average temperature of the warmer month at 22°C (Junges and Fontana, 2011). The region is located in the Plateau North-Rio-Grandense, formed by basaltic rocks of the Mesozoic Era. Altitudes decrease in the East-West direction, varying from 1300 meters to 700 meters (Becker and Nunes, 2012). The vegetation is formed by grasslands, Subtropical Forest and Araucária Forest (the name refers to the predominant species, *Araucária angustifolia*) (Moreira 2007). This region has a land use configuration formed by large concentration of small grain-producing properties.

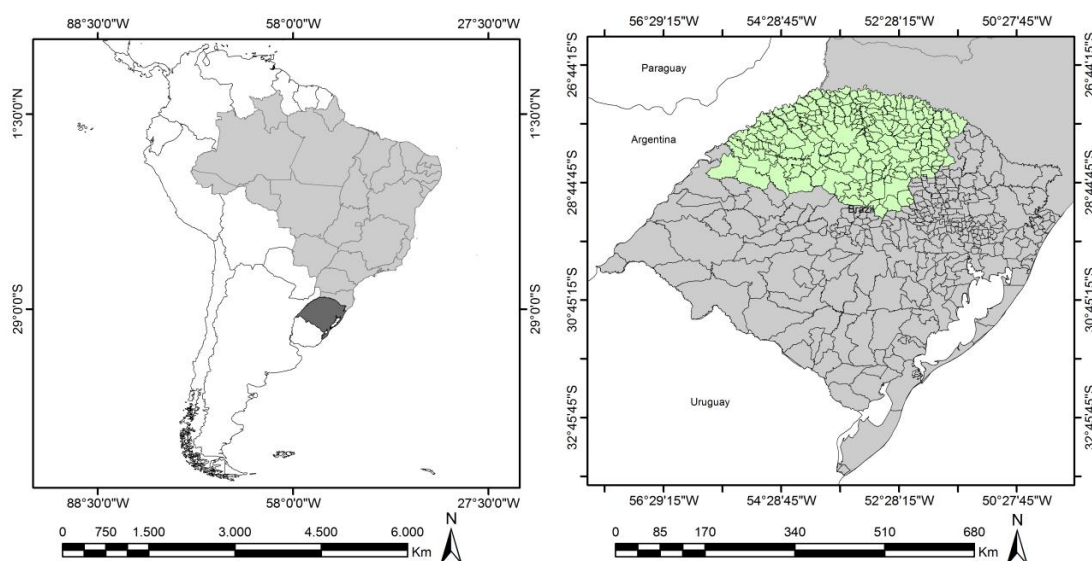


Figure 3. 1. Location map of the study area.

3.3. MATERIAL AND METHODS

3.3.1. MODIS data

The research used the MODIS Surface Reflectance product (MOD09Q1), with a spatial resolution of 250 meters. This product is an 8-day composition, which selects the best pixels with low view angle, the absence of clouds or cloud shadow, and aerosol loading (Gumma et al. 2015). The time series used corresponds to h13v11 tile downloaded from the National Aeronautics and Space Administration (NASA) (<http://reverb.echo.nasa.gov>). The MODIS Reprojection Tool (MRT) was used to convert the images in the sinusoidal projection and Hierarchical Data Format (HDF) format to the geographic coordinate system (WGS84 datum). The images correspond to the vegetative cycle of the wheat in the State of Rio Grande do Sul, when the plantation occurs during May, June and July and the harvest during October, November, December, crop year 2013.

3.3.2. Elaboration and filtering of the NDVI temporal cube

The Normalized Difference Vegetation Index (NDVI) calculation uses the red (ρ_{RED}) (620-670 nm) and near-infrared (ρ_{NIR}) (841-875 nm) bands in the following equation (Rouse et al. 1973):

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (1)$$

Subsequently, the images of the time series were grouped in a cube, where the "x" and "y" axes correspond to the geographic coordinates and the "z" axis to the increasingly ordered NDVI temporal signatures (Carvalho et al. 2008).

The time series are susceptible to noise from data acquisition errors, the presence of clouds and shadows, and atmospheric inferences (Hird and McDermid 2009; Carvalho Junior et al. 2013). In this research, we use the Savitzky-Golay (S-G) smoothing method to minimize noise (Savitzky and Golay 1964), which is used in the various signal processing such as analytical chemistry, geophysical (e.g., Baba et al. 2014), electrocardiograph (e.g., Birle et al. 2015), among others. This convolution technique fits an array of "n" equally-spaced points to a polynomial, which is a subgroup of complete temporal series. The S-G method tends to preserve peak attributes such as height, width, asymmetry, maximum and minimum point (Schefer 2011). This peak shape conservation property by the S-G filters is desirable in the remote sensing time-series processing for vegetation analysis, which uses these attributes to distinguish the phenological phases (Chen, et al. 2004; Abade et al. 2015). This explains the wide use of the S-G method for image denoising in several applications such as to detect natural vegetation (e.g., Fontana et al. 2008; Hüttich et al. 2009), phenological analysis of crops (Ren et al. 2008; Pan et al. 2015), fire events (e.g., Veraverbeke et al. 2010; Carvalho Júnior et al. 2015), among others. **Figure 3.2** demonstrates the effect of the S-G filtering for a wheat temporal profile present in the study area.

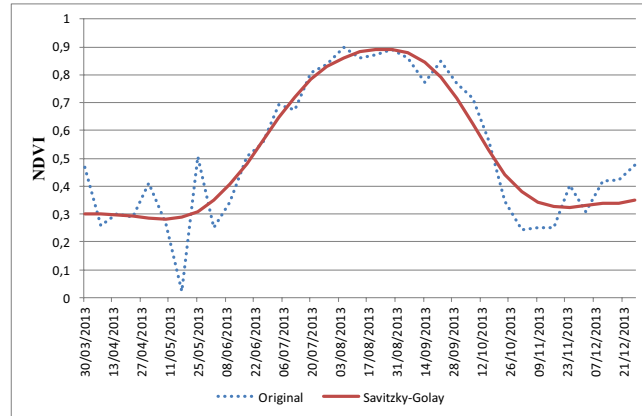


Figure 3. 2. Filtering of the MODIS NDVI time series by the Savitzky-Golay method.

3.3.3. Selection and classification of temporal signatures

The selection of NDVI temporal signatures of wheat considered a spatially distributed sampling, depicting the different sowing dates and vegetation growth. Thus, 13 temporal signatures were collected that exemplify the different phenological patterns of wheat in the region. The classification procedure considered two problems: (a) variability of the temporal signatures of the wheat crop due to the different dates of planting and phenological variations; and (b) the presence of large numbers of targets besides wheat crops (natural vegetation, other crops, built areas, among others), which are included in traditional supervised classifications to improve the performance of the mapped object. In this context, this study aims to develop a wheat crop detection, considering its variability and does not require sampling of the other elements present in the scene. Thus, the formulated method is based on two steps: (a) establishing a metric that highlights the area of interest; and (b) apply a mask from the threshold value. In this approach, searches for several targets have been developed from the threshold value determination for spectral indices: (a) land-cover change detection from NDVI differencing (Lunetta et al. 2006); (b) burned area detection using NDVI and Normalized Burned Ratio (NBR) differencing (Escuin et al. 2008), and (c) detection of built-up areas using the built-up index-based index (IBI) (Xu et al. 2008).

In this research, we adopted the nearest neighbor metrics from the k-NN algorithm to evidence wheat growing areas, instead of taking a spectral index. The k-NN algorithm is a non-parametric multivariate method and one of the simplest algorithms among machine learning algorithms, but yields highly competitive results (Cover and Hart, 1967; Fix and Hodges, 1951). The classification criterion of a point is the majority class of its k neighbors, where k is a user-defined constant consisting of a positive integer. If k is equal to 1, the method is called the "nearest neighbor". The adoption of this variant of the k-NN method was due to the quality of the selected samples and the variation of their signatures. This method allows "learning by example" since it uses a known response close to the given vector.

Typically, the metric used to identify nearest neighbors is the Euclidean distance. In addition to this metric, we evaluated the measure of cosine similarity (Salton and McGill, 1983), which is widely used in spectral classification from the Spectral Angle Mapper algorithm (Kruse et al. 1993). This metric determines the spectral similarity between two vectors, calculating the angle between them in a space with dimensionality equal to the number of bands. The value similarity is maximum when the vectors are in the same direction (angle equal to 0°) and becomes minimum when perpendicular (angle equal to 90°). The inclusion of cosine similarity in the calculation of the k-NN method was performed in the *Abilius* program (Carvalho Júnior et al. 2011) using C ++ language. The k-NN method is

widely used to classify remote sensing data, mainly in forestry studies (e.g., Chirici et al. 2016; Franco-Lopez et al. 2001; Magnussen et al. 2009).

The detection of the best threshold value between the wheat and non-wheat areas in the nearest neighbor metric images adopted the method proposed by Carvalho Júnior et al. (2015). Initially, this algorithm was developed to detect the best threshold value between burned and unburned areas in NBR images. The method considers two input data: (a) reference map; and (b) metric image. The technique automatically applies a series of threshold values and compares the resulting classifications with a reference map from the Kappa index. The highest Kappa index is the best threshold value. The reference map was elaborated from the visual interpretation of Landsat 8 images in an area corresponding to 5% of the total area studied.

3.3.4. Accuracy analysis

In the accuracy analysis, we used the confusion matrix, Overall coefficient, Kappa index and errors of commission and omission (Congalton and Green, 1993). The accuracy indices were obtained from the comparison of the classified images and the ground truth image produced from the visual interpretation of Landsat-8 images and field data, corresponding to 10% of the total area mapped. The location of this ground truth map was different from the reference map for the detection of the best threshold value in the wheat classification.

McNemar's test was used in the evaluation of the statistical significance between the two nearest neighbor classifications (McNemar, 1947). This test is a non-parametric method that establishes a binary distinction between correct and incorrect classes from the same set of samples, assuming a confusion matrix with a size of 2 by 2 (Foody, 2004). The chi-square (χ^2) statistic with one degree of freedom is computed from confusion matrix by the following equation:

$$\chi^2 = \frac{(f_{12} - f_{21})^2}{f_{12} + f_{21}}$$

Where f_{12} is the number of wrong classifications by method 1 but correct by method 2, and f_{21} is the number of correct classifications by method 1 but wrongly classified by method 2. This accuracy comparison based on related samples are deep-rooted in the literature, containing broad application (e.g., Leeuw et al. 2006, Manandhar et al. 2009). In the McNemar's test, we used 1000 samples.

Moreover, we compared the planted area by time series classification with the official statistics from the Brazilian Institute of Geography and Statistics (IBGE) obtained by direct interviews. In this approach, several studies compare agricultural census statistics and maps from digital image processing (Biradar and Xiao, 2011; Frolking et al. 1999, Junges and Fontana, 2009; Qiu et al. 2003; Zhang et al. 2011).

3.4. RESULTS

3.4.1. Analysis of temporal signatures of wheat cultivation

Figure 3.3 shows the variation of wheat temporal signatures with different dates to start planting and maximum NDVI position since each region has specific differences of the agricultural calendar. The minimum values of NDVI occurred in April in the period before sowing, and maximum values occurred from July during the vegetative development stage of the plant. The decrease of the NDVI value from September is a function of leaf senescence

when the plants began to reach maturity. Junges and Fontana (2009) found similar results for the temporal signatures of winter crops in the Rio Grande do Sul.

3.4.2. Results of the ratings of the time series

Figures 3.4 represents the Kappa index curves in the detection of the best threshold value for wheat crop classification, considering the nearest neighbor metrics of the Euclidean distance and cosine similarity. The optimal threshold for Euclidean distance was 0.710 with Kappa index of 78.15 (**Figure 3.5a**), while for cosine similarity was 0.203 degrees with Kappa index of 75.60 (**Figure 3.5b**).

Figure 3.5 shows the wheat planting masks from the two metrics. The classification using the Euclidean distance metric showed an overall accuracy of 90.01%, Kappa index of 0.75, an error of omission of 21.17%, and an error of commission of 13.06% (**Table 3.1**). This performance was slightly higher than that of the cosine similarity metric with an overall accuracy of 89.32%, Kappa index of 0.74, an error of omission of 21.50%, and an error of commission of 14.93%.

The McNemar's test certified the resemblance of the nearest neighbor classifications using Euclidean distance and cosine similarity (**Table 3.2**). The calculated χ^2 was 1.125 so the test provides evidence to accept the null hypothesis of marginal homogeneity states or equal classifier performance at the 5% level of significance ($p = 0.05$). In addition, the 2x2 contingency table demonstrated 96.8% of equal pairs (correct vs. correct and incorrect vs. incorrect).

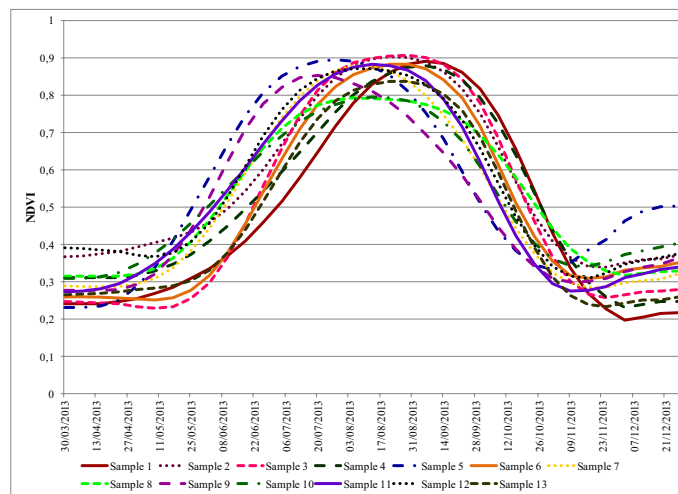


Figure 3. 3. MODIS NDVI temporal signatures for wheat crop in the Northwest Rio Grande do Sul.

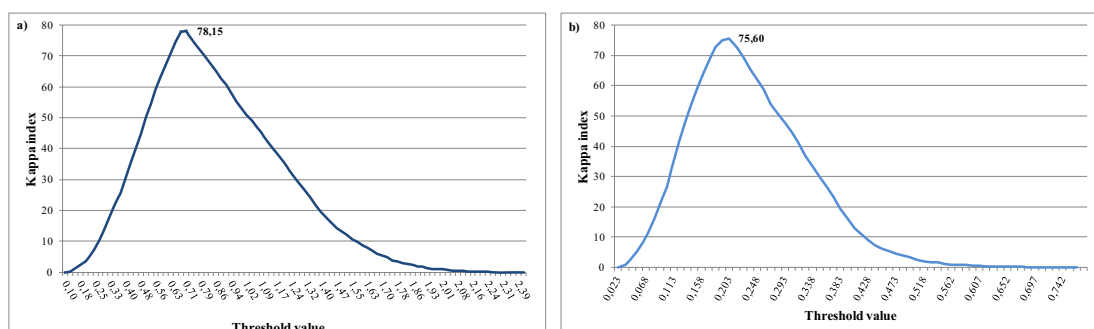


Figure 3. 4. Kappa index curve for the definition of threshold values for the two metrics of the nearest neighbor: (a) Euclidean distance and (b) cosine similarity.

Table 3. 1. Confusion Matrix of the classifications based on the nearest neighbor metric images: Euclidean distance and cosine similarity. Where "EO" is the error of omission and "EC" is the error of commission.

Euclidian Distance Overall Accuracy: 90.01 Kappa Coefficient: 0.75				Cosine similarity Overall Accuracy: 89.32 Kappa Coefficient: 0.74			
	Wheat	No-Wheat	Total	C	Wheat	No-Wheat	Total
Wheat	78.83	5.14	27.43	Wheat	78.50	5.97	27.91
No-Wheat	21.17	94.86	72.57	No-Wheat	21.50	94.03	72.09
Total	100	100	100	Total	100	100	100
EO	21.17	5.14	-	E O	21.50	5.97	-
EC	13.06	8.82	-	E C	14.93	9.02	-

Table 3. 2. McNemar's test between the two classifications by the nearest neighbor using Euclidean distance and the cosine similarity.

Allocation	Euclidian Distance		
	Correct	Incorrect	Total
Cosine Similarity			
Correct	881	13	894
Incorrect	19	87	106
Total	900	100	1000
χ^2	1.125		

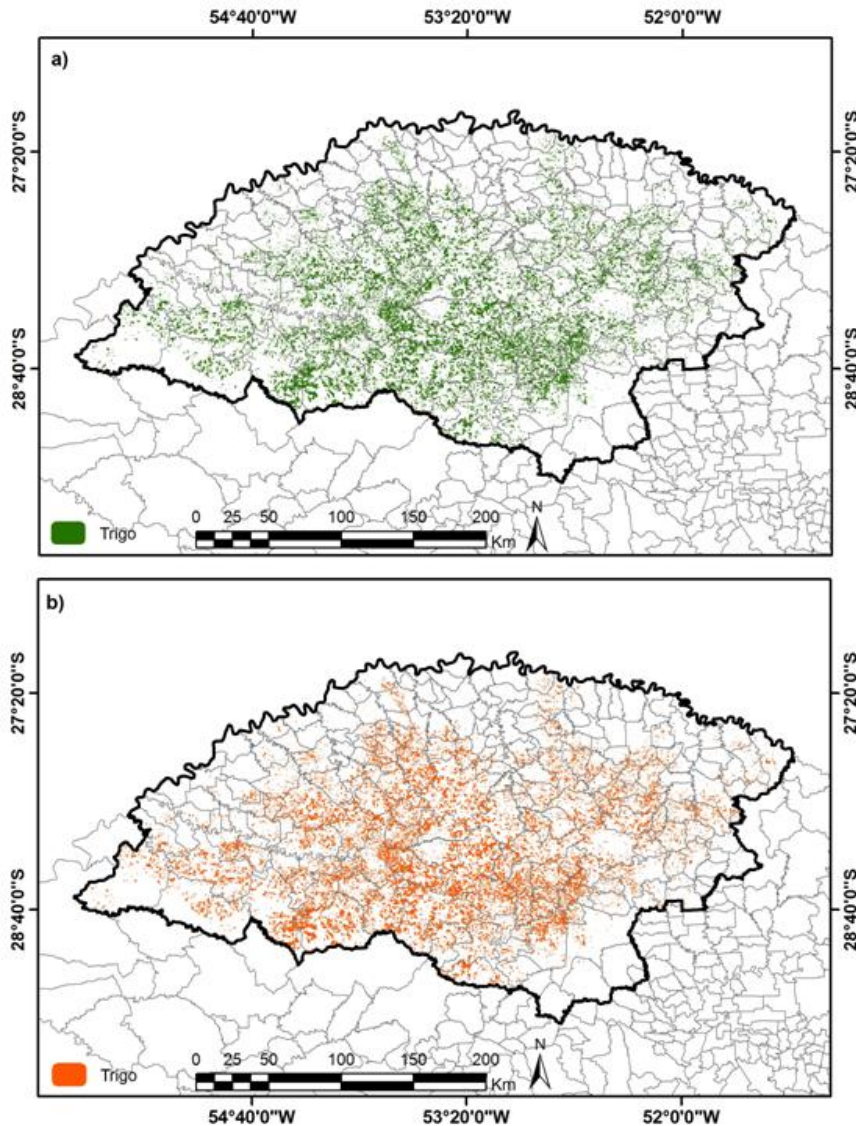


Figure 3. 5. Wheat planting map by using the threshold value in the nearest neighbor metric images: (a) Euclidian distance, (b) cosine similarity.

3.5. DISCUSSION

The results demonstrate a significant variation of the wheat temporal curves. Different factors caused the phenological variation of the cultivated areas: genetic varieties, latitude positioning, temperature, precipitation, photoperiod, soil altimetry, soil, plant age, pest presence, among others. In addition, the temporal behavior of the crop is influenced by farmer decisions and management practices (Wardlow and Egbert 2008). The results may also present changes due to space-time resolution, smooth procedures, and methods for extracting phenological features (Chu et al. 2016). Therefore, the variability of temporal signatures can be high even within the same culture.

Specifically, in the analysis of wheat phenological profiles using remotely-sensed temporal series, several studies have shown spatial variations. Huang et al. (2016) found relationships between the different growth stages of winter wheat and meteorological factors. Franch et al. (2015) verified temporal variations of the NDVI peak position due to climatic variations. Lu et al. (2014) evidenced spatial changes in wheat phenology in relation to the

degree of latitude in the Northern China Plain. Mafron et al. (2017) demonstrated a marked interannual variability in sowing times for winter wheat mainly due to the type of the previous crop and the climatic conditions during the summer period before to planting. These variations explain the limitation of establishing a single threshold value on a vegetation index in order to detect vegetation types and crops (White et al. 1997; Fisher et al. 2006). Therefore, many studies of wheat crop detection adopt methods based on specific phenological metrics (Lobell et al. 2013; Vyas et al. 2013; Chu et al. 2016).

The proposed method does not need to extract phenological metrics for classification. The classification is based on K-NN method metric (distance or similarity) that can consider training samples with different signatures of the same target, and the result depends only on those that are important for each situation. This approach makes the method more flexible and suitable for studies of vegetation and crops that describe marked spatiotemporal variations. In the case of the nearest neighbor, only a single temporal signature within the sampling set is used for the metric calculation. The metric values acquire similar values from the distinct temporal signatures within the sample set, considering different sowing dates or amplitude variations. One limitation of this method is the need to have a sample set that characterizes the present variations of the intended target. In the k-NN method, the study evaluated the metrics of Euclidean distance and cosine similarity. The results demonstrated have a slight advantage of the Euclidian distance in relation to cosine similarity.

The wheat mapping showed accuracy values compatible with other surveys with the MODIS sensor and obtained a planted area close to the estimated values by traditional agricultural surveys. In the analysis of accuracy, an error source frequently reported in agricultural crop mappings using MODIS NDVI time series is the use of higher resolution ground truth maps (e.g., Li and Tiam, 2011; Junges et al. 2013; Dempewolf et al. 2014). This problem is minimized in extensive and continuous plantation mappings (Wardlow et al. 2006; Wardlow and Egbert, 2008), while the errors are aggravated in small areas of cultivation, for example, small wheat farms in the region Yellow River Delta, China (Chu et al. 2016). In the study area, wheat cultivation occurs in small farms, resulting in an overestimation of the planted area and a high error concentration at the property boundaries. Therefore, part of the error is inherent in the accuracy analysis that compares data with different resolution. The use of new sensors with high spatial and temporal resolution allows minimizing these limitations in the precision of the crop areas.

3.6. CONCLUSION

This research developed a method to classify wheat crops using MODIS NDVI time series from the threshold values of distance and similarity metrics of the nearest neighbor. This approach allows similar results for different behaviors of wheat phenology, which presents a significant variation for the study area. Besides, the method does not require the extraction of phenological metrics or the sampling of other targets present in the scene. One proposed innovation was the use of the cosine similarity metric for the calculation of the nearest neighbor. This similarity metric is used in the SAM classifier, widely applied in spectral classification. In the wheat detection, the nearest neighbor methods using Euclidean distance and cosine similarity obtained equivalent results according to McNemar's test. The results of wheat planted area are compatible with the agricultural census using traditional field surveys in the same year of analysis. Government departments committed to the dissemination of statistics on agricultural production may use the methodology to assist in the work of calculating planted areas.

REFERENCES

- Abade NA, de Carvalho Júnior OA, Guimarães RF, Oliveira SN. 2015. Comparative analysis of MODIS time-series classification using support vector machines and methods based upon distance and similarity measures in the Brazilian Cerrado-Caatinga boundary. *Remote Sensing*, 7(9):12160-12191.
- Asseng S, Ewert F, Martre P, Rötter R, Lobell D, Cammarano D, Kimball BA, Ottman M, Wall G, White J, et al. 2014. Rising temperatures reduce global wheat production. *Nat. Clim. Change* 5:143–147.
- Atchison J, Head L, Gates A. 2010. Wheat as food, wheat as industrial substance; comparative geographies of transformation and mobility. *Geoforum*, 41(2):236-246.
- Baba K, Bahi L, Ouadif L. 2014. Enhancing geophysical signals through the use of Savitzky-Golay filtering method. *Geofísica Internacional*, 53(4):399-409.
- Becker ELS, Nunes MP. 2012. Relevo do Rio Grande do Sul, e sua representação em maquete [Relief of Rio Grande do Sul, Brazil, and their representation in mockup]. *Revista Percorso* 4:113-132,
- Becker-Reshef I, Vermote E, Lindeman M, Justice C. 2010. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment*, 114(6):1312-1323.
- Biggs TW, Thenkabail PS, Gumma MK., Scott CA, Parthasaradhi GR, Turrall, HN. 2006. Irrigated area mapping in heterogeneous landscapes with MODIS time series, ground truth and census data, Krishna Basin, India. *International Journal of Remote Sensing* 27(19):4245-4266.
- Biradar CM, Xiao X. 2011. Quantifying the area and spatial distribution of double-and triple-cropping croplands in India with multi-temporal MODIS imagery in 2005. *International Journal of Remote Sensing* 32(2):367-386.
- Birle A, Malviya S, Mittal D. 2015. Noise removal in ECG signal using Savitzky-Golay filter. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)* 4(5):1331-1333.
- Camponogara A, Gallio E, Borba WF, Georgin J. 2015. O atual contexto da produção de trigo no Rio Grande do Sul [The current context of wheat producing in Rio Grande do Sul]. *Revista Eletrônica em Gestão, Educação e Tecnologia Ambiental* 19(2):246-257.
- Canziani JR, Guimarães VDA. 2009. O trigo no Brasil e no mundo: cadeia de produção, transformação e comercialização [Wheat in Brazil and in the world: production, processing and marketing chain]. In: Cunha GR, editor, *Oficina sobre trigo no Brasil: bases para a construção de uma nova triticultura [Workshop on wheat in Brazil: bases for the construction of a new Brazilian wheat crop]*. Passo Fundo, RS, Brazil: Embrapa Trigo, p. 29-72.
- Carvalho Júnior OA, Guimarães RF, Gillespie AR, Silva NC, Gomes RAT. 2011. A New approach to change vector analysis using distance and similarity measures. *Remote Sensing* 3: 2473-2493.
- Carvalho Júnior OA, Guimarães RF, Silva CR, Gomes RAT. 2015. Standardized time-series and interannual phenological deviation: new techniques for burned-area detection using long-term MODIS-NBR Dataset. *Remote Sensing* 7: 6950-6985.
- Carvalho OA, Guimarães RF, Silva NC, Gillespie AR, Gomes RAT, Silva CR, de Carvalho APF. 2013. Radiometric normalization of temporal images combining automatic detection of

pseudo-invariant features from the distance and similarity spectral measures, density scatterplot analysis, and robust regression. *Remote Sensing*, 5(6):2763-2794.

Carvalho Júnior OA, Sampaio CDS, Silva NCD, Couto Júnior AF, Gomes RAT, Carvalho APF, Shimabukuro YE. 2008. Classificação de padrões de savana usando assinaturas temporais NDVI do sensor MODIS no Parque Nacional Chapada dos Veadeiros. *Revista Brasileira de Geofísica* 26(4):505-517.

Chen J, Jönsson P, Tamura M, Gu Z, Matsushita B, Eklundh L. 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment* 91: 332–344.

Chirici G, Mura M, Mcinerney D, Py N, Tomppo EO, Waser LT, Travaglini D, Mcroberts RE. 2016. A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. *Remote Sensing of Environment* 176: 282-294.

Chu L, Liu QS, Huang C, Liu GH. 2016. Monitoring of winter wheat distribution and phenological phases based on MODIS time-series: A case study in the Yellow River Delta, China. *Journal of Integrative Agriculture* 15: 2403-2416.

Congalton RG, Green K. 1993. A Practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering & Remote Sensing* 59(5):641-644.

Cover T, Hart P. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13(1):21-27.

Cunha GR, Haas JC, Maluf JRT, Caramori PH, Assad ED, Braga HJ, Zullo Júnior J, Lazzarotto C, Gonçalves S, Wrege M, Brunetta D, Dotto SR, Pinto HS, Brunini O, Thomé VMR, Zampieri SL, Pasinato A, Pimentel MBM, Pandolfo C. 2001. Agricultural zoning and sowing dates for wheat in Brazil. *Revista Brasileira de Agrometeorologia*, 9(3):446-459.

Dempewolf J, Adusei B, Becker-Reshef I, Hansen M, Potapov P, Khan A, Barker B. 2014. Wheat yield forecasting for Punjab Province from vegetation index time series and historic crop statistics. *Remote Sensing* 6(10):9653-9675.

Dixon J, Braun HJ, Crouch J. 2009. Overview: transitioning wheat research to meet the future demands of the developing world. In: Dixon J, Braun HJ, Kosina P, Crouch J., editors. *Wheat facts and futures 2009*. Mexico, DF: International Maize and Wheat Improvement Center (CIMMYT). Chapter 1, p. 1-25.

Dixon J, Nalley L, Kosina P, La Rovere R, Hellin J, Aquino P. 2006. Adoption and economic impact of improved wheat varieties in the developing world. *Journal of Agricultural Science*, 144, 489–502.

Dong T, Liu J, Qian B, Zhao T, Jing Q, Geng X, Wang J, Huffman T, Shang J. 2016. Estimating winter wheat biomass by assimilating leaf area index derived from fusion of Landsat-8 and MODIS data. *International Journal of Applied Earth Observation and Geoinformation*, 49, 63-74.

Escuin S, Navarro R, Fernandez P. 2008. Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *International Journal of Remote Sensing* 29(4):1053-1073.

Feng M, Yang W, Zhang D, Cao L, Wang H, Wang Q. 2009. Monitoring planting area and growth situation of irrigation-land and dry-land winter wheat based on TM and MODIS data. *Transactions of the Chinese Society of Agricultural Engineering*, 25(3):103-109.

- Feng MC, Yang WD, Cao LL, Ding GW. 2009. Monitoring winter wheat freeze injury using multi-temporal MODIS data. *Agricultural Sciences in China*, 8(9):1053-1062.
- Fisher JJ, Mustard JF, Vadeboncouer MA. 2006. Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sensing of Environment* 100: 265–279.
- Fix E, Hodges JL. 1951. Discriminatory analysis — Nonparametric discrimination: Consistency properties. *International Statistical Review* 57: 238–247.
- Fontana F, Rixen C, Jonas T, Aberegg G, Wunderle S. 2008. Alpine grassland phenology as seen in AVHRR, VEGETATION, and MODIS NDVI time series-A comparison with in situ measurements. *Sensors* 8(4):2833-2853.
- Foody GM. 2004. Thematic map comparison. *Photogrammetric Engineering & Remote Sensing*, 70(5):627-633.
- Franch B, Vermote EF, Becker-Reshef I, Claverie M, Huang J, Zhang J, Justice C, Sobrinho JA. 2015. Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information. *Remote Sensing of Environment* 161: 131–148.
- Franco-Lopez H, Ek AR, Bauer ME. 2001. Estimation and mapping of forest stand density: volume and cover type using the k-nearest neighbors method. *Remote Sensing of the Environment* 77: 251 – 274.
- Frolking S, Xiao X, Zhuang Y, Salas W, Li C. 1999. Agricultural land-use in China: A comparison of area estimates from ground-based census and satellite-borne remote sensing. *Global Ecology and Biogeography* 8(5):407-416.
- Gumma MK, Mohanty S, Andrew N, Rala A, Irshad AM, Das SR. 2015. Remote sensing based change analysis of rice environments in Odisha, India. *Journal of Environmental Management* 148: 31-41.
- Hao P, Wang L, Zhan Y, Niu Z. Using Moderate-Resolution Temporal NDVI profiles for high-resolution crop mapping in years of absent ground reference data: a case study of Bole and Manas Counties in Xinjiang, China. *ISPRS International Journal of Geo-Information* 5(67):1-23, 2016.
- Hird JN, McDermid GJ. 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment*, 113(1):248-258.
- Huang Q, Wang LM, Chen ZX, Hang L. 2016. Effects of meteorological factors on different grades of winter wheat growth in the Huang-Huai-Hai Plain, China. *Journal of Integrative Agriculture* 15(11):2647-2657.
- Hüttich C, Gessner U, Herold M, Strohbach BJ, Schmidt M, Keil M, Dech S. 2009. On the suitability of MODIS time series metrics to map vegetation types in dry savanna ecosystems: A case study in the Kalahari of NE Namibia. *Remote Sensing* 1(4):620-643.
- [IBGE] Instituto Brasileiro de Geografia e Estatística. 2016. *Levantamento Sistemático da produção Agrícola: pesquisa mensal de previsão e acompanhamento das safras agrícolas no ano civil [Systematic survey of agricultural production: monthly survey of the forecast and monitoring of agricultural crops in the calendar year]*. Rio de Janeiro: Fundação Instituto Brasileiro de Geografia e Estatística, 29(5):1-79.
- Junges AH, Fontana DC, Pinto DG. 2013. Identification of croplands of winter cereals in Rio Grande do Sul state, Brazil, through unsupervised classification of normalized difference vegetation index images. *Engenharia Agrícola* 33(4):883-895.

- Junges AH, Fontana DC. 2009. Desenvolvimento das culturas de cereais de inverno no Rio Grande do Sul por meio de perfis temporais do índice de vegetação por diferença normalizada [Evaluation of winter cereal crop development in Rio Grande do Sul, Brazil, throughout the temporal profiles of normalized difference vegetation index]. *Ciência Rural* 39: 1349-1355,
- Junges AH, Fontana DC. 2011. Modelo agrometeorológico-espectral de estimativa de rendimento de grãos de trigo no Rio Grande do Sul. *Revista Ceres* 58: 9-16,
- Justice CO, Vermote E, Townshend JR, Defries R, Roy DP, Hall DK, Salomonson VV, Privette JL, Riggs G, Strahler A, et al. 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1228-1249.
- Kouadio L, Duveillerb G, Djabya B, Jarroudia ME, Defournyb B. 2012. Estimating regional wheat yield from the shape of decreasing curves of green area index temporal profiles retrieved from MODIS data. *International Journal of Applied Earth Observation and Geoinformation* 18: 111–118.
- Kruse FA, Lefkoff AB, Boardman JB, Heidebrecht KB, Shapiro AT, Barloon PJ, Goetz AFH. 1993. The Spectral Image Processing System (SIPS) - interactive visualization and analysis of imaging spectrometer Data. *Remote Sensing of the Environment* 44: 145–163.
- Leeuw J, Jia H, Yang L, Liu X, Schmidt K, Skidmore AK. 2006. Comparing accuracy assessments to infer superiority of image classification methods. *International Journal of Remote Sensing*, 27(1):223-232.
- Li S, Tian Q. 2011. Winter wheat area extraction and estimation based on MODIS-VI time series and multi-temporal HJ CCD images in Jiangsu province, China. In: Proceedings of the 2011 19th International Conference on Geoinformatics; Jun 26-24; Shanghai; IEEE. p. 1-6.
- Lobell D, Burke M. 2010. Economic impacts of climate change in agriculture. In: Reynolds MP, editor, *Climate change and crop production*. Wallingford: CABI.
- Lobell DB, Asner, GP. 2004. Cropland distributions from temporal unmixing of MODIS data. *Remote Sensing of Environment*, 93(3):412-422.
- Lobell DB, Ortiz-Monasterio JI, Sibley AM, Sohu VS. 2013. Satellite detection of earlier wheat sowing in India and implications for yield trends. *Agricultural Systems* 115: 137–143.
- Lu L, Wang C, Guo H, Li Q. 2014. Detecting winter wheat phenology with SPOT-VEGETATION data in the North China Plain. *Geocarto International* 29:244–255.
- Lunetta RS, Knight JF, Ediriwickrema, J, Lyon, JG, Worthy, LD. 2006. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment* 105(2):142-154.
- Magnussen S, Mcroberts RE, Tomppo EO. 2009. Model-based mean square error estimators for k-nearest neighbour predictions and applications using remotely sensed data for forest inventories. *Remote Sensing of Environment* 113: 476–488.
- Manandhar R, Odeh IO, Ancev T. 2009. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. *Remote Sensing* 1(3):330-344.
- Manfron G, Delmotte S, Busetto L, Hossard L, Ranghetti L, Brivio PA, Boschetti M. 2017. Estimating inter-annual variability in winter wheat sowing dates from satellite time series in Camargue, France. *International Journal of Applied Earth Observation and Geoinformation* 57: 190-201.

- McNemar Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157.
- Moreira I. O espaço rio-grandense [The space of the rio grande do sul]. São Paulo: Editora Ática, 2007. p. 96.
- Ozdogan M. 2010. The spatial distribution of crop types from MODIS data: temporal unmixing using independent component analysis. *Remote Sensing of Environment* 114: 1190–1204.
- Pan Y, Li L, Zhang J, Liang S, Zhu X, Sulla-Menashe D. 2012. Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. *Remote Sensing Environmental* 119: 232–242.
- Pan Z, Huang J, Zhou Q, Wang L, Cheng Y, Zhang H, Blackburn GA, Yan J., Liu J. 2015. Mapping crop phenology using NDVI time-series derived from HJ-1A/B data. *International Journal of Applied Earth Observation and Geoinformation* 34: 188–197.
- Pingali P. 2007. Westernization of Asian diets and the transformation of food systems: implications for research and policy. *Food Policy*, 32, 281–298.
- Qiu J, Tang H, Frolking S, Boles S, Li C, Xiao X, Liu J, Zhuang Y, Qin X. 2003. Mapping Single-, Double-, and Triple-crop Agriculture in China at $0.5^{\circ} \times 0.5^{\circ}$ by Combining County-scale Census Data with a Remote Sensing-derived Land Cover Map. *Geocarto International* 18(2):3–13.
- Ren J, Chen Z, Zhou Q, Tang H. 2008. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. *International Journal of Applied Earth Observation and Geoinformation* 10(4):403–413.
- Rouse JW, Haas RH, Schell JA, Deering DW. 1973. Monitoring vegetation systems in the Great Plains with ERTS. In: Third Earth Resources Technology Satellite-1 Symposium, Greenbelt, 1973, *Proceedings NASA SP-351*, p. 301–317.
- Salton G, McGill MJ. 1983. *Introduction to modern information*. Philadelphia, PA: American Association for Artificial Intelligence Retrieval.
- Savitzky A, Golay MJE. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* 36(8):1627–1639.
- Schefer RW. 2011. What is a Savitzky-Golay filter? *IEEE Signal Process. Mag.* 28, 111–117.
- Shiferaw B, Smale M, Braun HJ, Duveiller E, Reynolds M, Muricho G. 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security*, 5(3):291–317.
- Silva RR, Benin G, da Silva GO, Marchioro VS, de Almeida, JL, Matei G. 2011. Adaptabilidade e estabilidade de cultivares de trigo em diferentes épocas de semeadura, no Paraná [Adaptability and stability of wheat cultivars at different sowing dates in the state of Paraná, Brazil]. *Pesquisa Agropecuária Brasileira*, 46(11):1439–1447.
- Sun H, Xu A, Lin H, Zhang L. 2012. Winter wheat mapping using temporal signatures of MODIS vegetation index data. *International Journal of Remote Sensing* 33(16):5026–5042.
- Veraverbeke S, Lhermitte S, Verstraeten WW, Goossens R. 2010. The temporal dimension of differenced Normalized Burn Ratio (dNBR) fire/burn severity studies: The case of the large 2007 Peloponnese wildfires in Greece. *Remote Sensing of Environment*, 114, 2548–2563.

- Vyas S, Nigam R, Patel NK, Panigrahy S. 2013. Extracting regional pattern of wheat sowing dates using multispectral and high temporal observations from indian geostationary satellite. *J. Indian Soc. Remote Sens.* 41: 855–864.
- Wardlow BD, Egbert SL. 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment* 112: 1096–1116.
- Wardlow BD, Kastens JH, Egbert SL. 2006. Using USDA crop progress data for the evaluation of greenup onset date calculated from MODIS 250-meter data. *Photogrammetric Engineering & Remote Sensing* 72(11):1225-1234.
- Wenbo X, Guoping Z, Jinlong F, Yonglan, Q. 2007. Remote sensing monitoring of winter wheat areas using MODIS data. *Transactions of the Chinese Society of Agricultural Engineering*, 2007(12).
- White MA, Thornton PE, Running SW. 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles* 11: 217–234.
- Wit A, Duveiller G, Defourny P. 2012. Estimating regional winter wheat yield with WOFOST through the assimilation of green area index retrieved from MODIS observations. *Agricultural and Forest Meteorology*, 164:39-52.
- Xiao X, Boles S, Frolking S, Li C, Babu JY, Salas W, Moore III B. 2006. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment* 100(1):95-113.
- Xu H. 2008. A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing* 29(14):4269-4276.
- Yan H, Fu Y, Xiao X, Huang HQ, He H, Ediger L. 2009. Modeling gross primary productivity for winter wheat–maize double cropping system using MODIS time series and CO₂ eddy flux tower data. *Agriculture, Ecosystems & Environment*, 129(4):391-400.
- Yi Y, Yang D, Huang J, Chen D. 2008. Evaluation of MODIS surface reflectance products for wheat leaf area index (LAI) retrieval. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(6):661-677.
- Zhang S, Lei Y, Wang L, Li H, Zhao H. 2011. Crop Classification Using MODIS NDVI Data Denoised by Wavelet: A Case Study in Hebei Plain, China. *Chinese Geographical Science* 21(3):322–333.

CAPÍTULO 4 – MAPEAMENTO DE ÁREAS COM CULTIVOS DE ARROZ NO ESTADO DE SANTA CATARINA, ANO SAFRA 2014/2015, UTILIZANDO A TÉCNICA DO VIZINHO MAIS PRÓXIMO (K-NN), NA SÉRIE TEMPORAL NDVI MODIS

RESUMO

O arroz é o segundo cereal mais cultivado no mundo, sendo a base alimentar de mais de 50% da população mundial. Monitorar o cultivo de arroz em escalas regionais e globais é de vital importância para equilibrar a produção e as necessidades alimentares, defendendo ao mesmo tempo o ambiente. O objetivo deste trabalho é identificar a cultura do arroz ano safra 2014/2015 no Estado de Santa Catarina usando o método de classificação do vizinho mais próximo (K-NN) a partir de duas métricas: similaridade e distância euclidiana. A metodologia para o mapeamento de arroz envolveu as seguintes fases: aquisição de imagens MODIS, processamento das imagens, seleção das assinaturas temporais, classificação baseada no vizinho mais próximo (um caso específico do método K-NN) das métricas de similaridade e distância e análise de acurácia. A classificação baseada no vizinho mais próximo estabeleceu duas etapas: geração de imagens métricas (distância Euclidiana e similaridade do cosseno) e definição do melhor valor de corte para distinguir a máscara de arroz. Diferentes perfis temporais de arroz foram encontrados devido ao calendário agrícola ser diversificado na região. A classificação do vizinho mais próximo (K-NN) mostrou-se viável para mapeamento de culturas de arroz. A distância Euclidiana apresentou um índice Kappa de 0,73, e demonstrou um melhor desempenho do que a semelhança do cosseno que apresentou um índice Kappa de 0,72, como foi comprovada pelo teste McNemar.

Palavras chaves: arroz, NDVI MODIS, K-NN, distância Euclidiana, semelhança do cosseno, fenologia

ABSTRACT

Rice is the second most cultivated cereal in the world, being the food base of more than 50% of the world population. Monitoring rice cultivation on regional and global scales is of vital importance in balancing production and food needs while defending the environment. The objective of this work is to identify rice crop year 2014/2015 in the State of Santa Catarina using the method of classification of the nearest neighbor (K-NN) from two metrics: similarity and Euclidean distance. The methodology for mapping rice involved the following phases: MODIS image acquisition, image processing, temporal signatures selection, classification based on the nearest neighbor (a specific case of the K-NN method) of similarity and distance metrics and analysis of accuracy. The classification based on the nearest neighbor established two steps: generation of metric images (Euclidean distance and cosine similarity) and definition of the best cut value to distinguish the rice mask. Different temporal profiles of rice were found due to the agricultural calendar being diversified in the region. The classification of the nearest neighbor (K-NN) proved to be feasible for mapping rice crops. The Euclidean distance presented a Kappa index of 0.73, and showed a better performance than the cosine similarity that presented a Kappa index of 0.72, as evidenced by the McNemar test.

Keywords: rice, NDVI MODIS, K-NN, Euclidian distance, cosine similarity, phenology

4.1. INTRODUÇÃO

A cultura do arroz ocupa mais de 11% das terras agrícolas do mundo e faz parte da dieta básica de aproximadamente 50% da população mundial (SUBUDHI *et al.*, 2006). Essa cultura agrícola ocupa a segunda posição de área colhida (154 milhões ha), perdendo apenas para o trigo (214 milhões ha) (MACLEAN *et al.*, 2002). O arroz é importante na segurança alimentar global, a produção destinada para o consumo humano representa 85% da produção total, comparado com 72% para o trigo e 19% para o milho (MACLEAN *et al.*, 2002). A Ásia produz mais de 90% do arroz do mundo, onde China e a Índia são os maiores produtores mundiais (MACLEAN *et al.*, 2002). O crescimento rápido da população tem pressionado o aumento da produção de arroz pela intensificação e expansão do cultivo para atender a

demanda crescente. Dessa forma, o monitoramento do cultivo de arroz em escala regional e global é de vital importância para equilibrar a produção com as necessidades alimentares e estabelecer estratégias para um crescimento sustentável (CHEN *et al.* 2012).

No Brasil, o arroz é produzido em todos os Estados (IBGE, 2016). A Região Sul do país é responsável por 63% da produção nacional, sendo que o Estado do Rio Grande do Sul é o maior produtor brasileiro com 54%. Dentre os demais Estados do Brasil destacam-se Mato Grosso e Maranhão com 9% e Santa Catarina com 7% (**Figura 4.1a e b**). Atualmente existem dois sistemas de cultivos de arroz no Brasil conforme sua geomorfologia: terras altas e várzeas (CONAB, 2015). O arroz de terras altas (arroz de sequeiro) é caracterizado pelo plantio em áreas não alagadas, sendo dependente das boas condições pluviométricas para o perfeito desempenho dos estádios fenológicos da cultura (CONAB, 2015). No sistema de várzea existem duas modalidades, cultivo com irrigação controlada e sem irrigação controlada. No primeiro, a cultura é irrigada por inundação contínua e controlada com a formação e manutenção de lâmina de água até a maturação do arroz. No segundo, o plantio é realizado em áreas de baixadas, nas quais as chuvas e as enchentes dos rios ou afloramento natural do lençol freático são as fontes de água para o desenvolvimento das plantas. Os Estados do Rio Grande do Sul e Santa Catarina usam a técnica de arroz irrigado, alcançando maior produtividade do que o cultivo de arroz sequeiro. No estado de Santa Catarina, 90% do sistema de cultivo utiliza sementes pré-germinadas, no qual a semeadura é realizada em lâmina de água. Uma das principais vantagens deste modo de produção é a elevação da produtividade (CONAB, 2015).

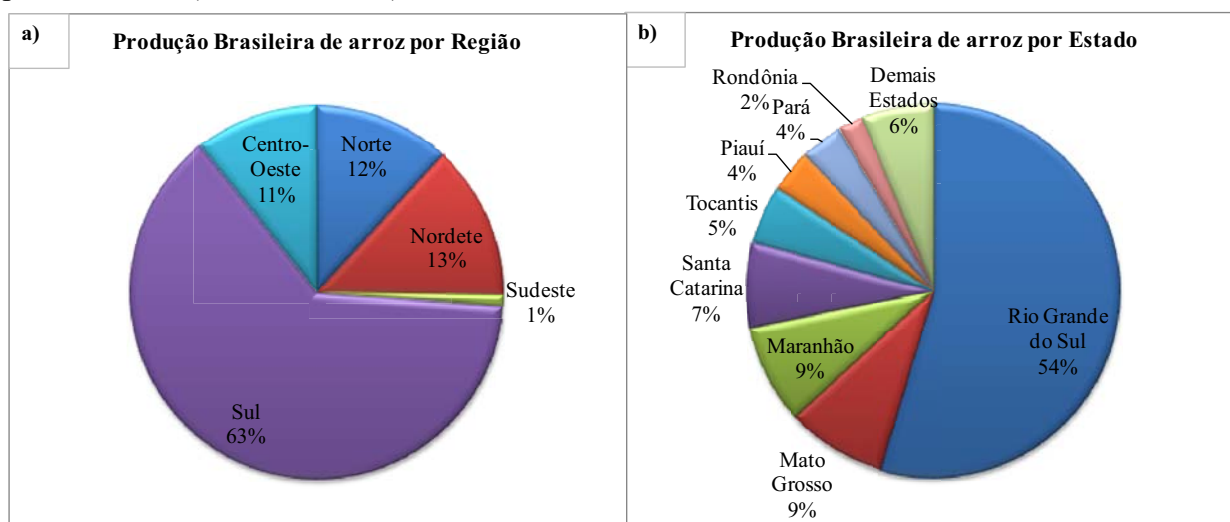


Figura 4. 1. Produção brasileira de arroz em área plantada por hectare. Fonte: IBGE, 2016.

A estimativa da produção agrícola brasileira das principais lavouras de grãos utiliza métodos baseados em entrevistas aos agricultores e entidades do setor agrícola. No entanto, a grande extensão do território brasileiro e a alta dinâmica espaço-temporal da atividade agrícola tornam aquisição dessas informações custosas e demoradas. Em compensação, a utilização do sensoriamento remoto e do Sistema de Informação Geográfica tem a capacidade de auxiliar na obtenção dessas informações de forma rápida, precisa e com menor custo (RIZZI & RUDORFF, 2005). Os mapeamentos da cultura do arroz com sensores remotos orbitais exploram o desenvolvimento fenológico da planta em seu ciclo de vida, sendo composto pelas fases de germinação, desenvolvimento vegetativo, floração, enchimento de grãos e maturação. Portanto, vários estudos de mapeamento das plantações de arroz têm utilizado dados temporais de sensores orbitais para demarcar as diferentes fases fenológicas, principalmente os dos satélites Landsat (JIN *et al.*, 2015; KONTGIS *et al.*, 2015; D'ARCO *et al.*, 2006) e do sensor Moderate Resolution Imaging Spectroradiometer (MODIS) devido à

sua alta resolução temporal (CHEN *et al.*, 2012; CLAUSS *et al.*, 2016; GUAN *et al.*, 2016; GUMMA *et al.*, 2014; PENG *et al.*, 2011; TELUGUNTLA *et al.*, 2015).

No Brasil, alguns estudos de séries temporais para detecção de arroz foram realizados no Estado do Rio Grande do Sul. Klering *et al.*, (2014), utilizaram imagens MODIS para definir as áreas cultivadas com arroz irrigado e monitorar o cultivo ao longo do ciclo para dez safras agrícolas. D'Arco *et al.*, (2006), estimaram a área plantada com arroz irrigado por meio de imagens dos sensores TM e ETM+ dos satélites Landsat 5 e 7, respectivamente.

O presente estudo possui como objetivo identificar a cultura do arroz a partir de séries temporais MODIS, usando o método de classificação do vizinho mais próximo (K-NN) a partir de duas métricas: similaridade e distância euclidiana, ano safra 2014/2015, no Estado de Santa Catarina.

4.2. ÁREA DE ESTUDO

A área de estudo localiza-se na porção leste do estado de Santa Catarina na Região Sul do Brasil (**Figura 4.2**). As áreas de plantio de arroz irrigado do Estado ocupam cinco regiões distintas por suas condições geográficas e edafoclimáticas: Alto, Médio/Baixo Vale do Itajaí, Litoral Norte e região Sul do Estado (SOSBAI, 2014). Santa Catarina possui a segunda maior produtividade nacional de arroz, mesmo com área bem inferior à de vários outros Estados. A mesorregião sul catarinense, formada pelas microrregiões de Araranguá, Tubarão e Criciúma, responde por 63,45% da área e 64,10% da produção estadual (safra 2015/2016). Sendo que, a microrregião de Araranguá é responsável por 35% da área e 35,35% da produção estadual (EPAGRI, 2016).

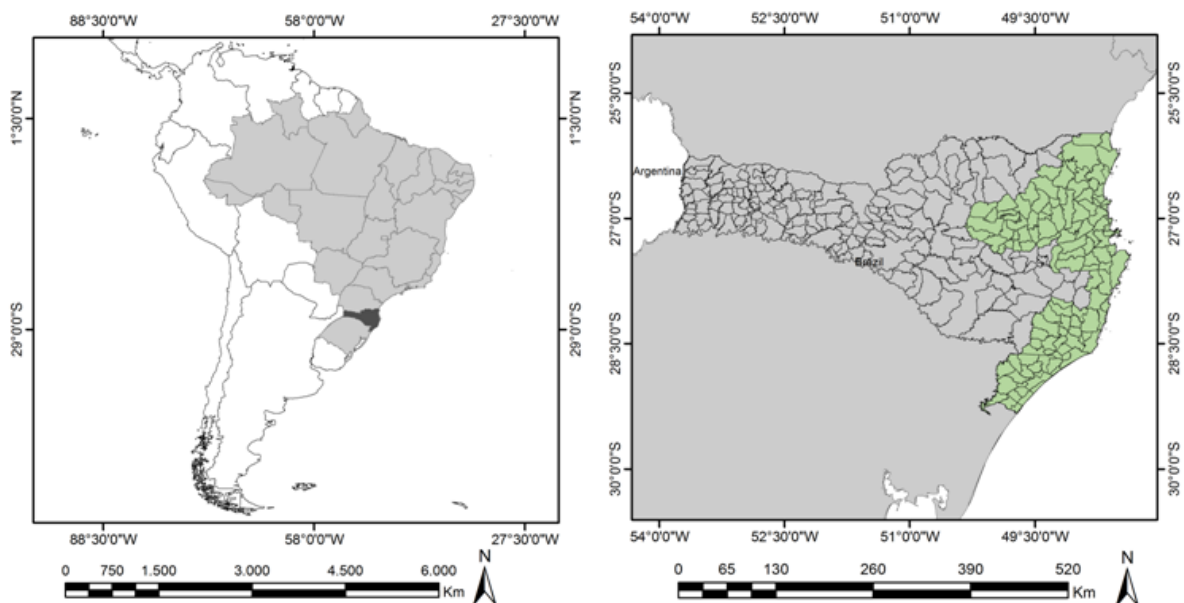


Figura 4. 2. Localização da área de estudo.

O clima de Santa Catarina é mesotérmico úmido (sem estação seca) (Cf) conforme a classificação de Köpen, incluindo dois subtipos: (a) Clima subtropical (Cfa), com temperatura média no mês mais frio inferior a 18°C (mesotérmico) e temperatura média no mês mais quente acima de 22°C, com verões quentes, geadas pouco frequentes e tendência de concentração das chuvas nos meses de verão; e (b) Clima temperado (Cfb) com temperatura média o mês mais frio abaixo de 18°C (mesotérmico), com verões frescos, temperatura média no mês mais quente abaixo de 22°C (PANDOLFO *et al.*, 2002). A geomorfologia do Estado

de Santa Catarina é formada pelas planícies costeiras, planalto das araucárias, escarpas e reversos da Serra do Mar e as serras do leste catarinense. A altitude varia de 10m (planícies costeiras) a 1.200m (Serra Geral). A vegetação é composta por vegetação florestal, campestre, mangues, formação de praias e dunas e restinga (POTTER *et al.*, 2004).

A cultura do arroz irrigado foi implantada pelos imigrantes italianos no começo do século XX no Vale do Itajaí, possivelmente, em virtude do ambiente com predomínio de solos argilosos mal drenados e inexistência de estação seca (CONAB, 2015). O arroz irrigado em Santa Catarina é caracterizado como uma cultura típica de pequena propriedade rural com mão de obra familiar, possuindo tamanho médio de 13,5 hectares (SOSBAI, 2014).

4.3. MATERIAL E METODOLOGIA

4.3.1. Dados MODIS

Nesse trabalho foi utilizado o produto de reflectância de superfície MOD09Q1, com resolução espacial de 250 metros e composição temporal de 8 dias, selecionado a partir da taxa de cobertura de observação, o ângulo de visão, a ausência de nuvens ou sombra de nuvens, e o carregamento de aerossol (GUMMA *et al.*, 2015). Segundo Guindin-Garcial *et al.*, (2012), a resolução temporal é relevante no monitoramento da fenologia das culturas agrícolas. Assim, a escolha das imagens de 8 dias apoia-se na capacidade de capturar todo o ciclo da cultura do arroz.

A série temporal utilizada corresponde ao tile h13v11 adquirida na página eletrônica da National Aeronautics and Space Administration (NASA) (<http://reverb.echo.nasa.gov>). As imagens disponibilizadas na projeção sinusoidal e no formato Hierarchical Data Format (HDF) foram convertidas para o sistema de coordenadas geográficas, datum Sistema Geodésico Mundial 84 (WGS84), por meio da ferramenta MODIS Reprojection Tool (MRT). As imagens usadas foram referentes ao período que engloba o ciclo vegetativo da cultura do arroz no Estado de Santa Catarina, conforme calendário agrícola da Conab (**Tabela 4.1**). A pesquisa utilizou 35 imagens para o ano safra 2014/2015, no período de agosto de 2014 a abril de 2015.

Tabela 4. 1. Calendário de plantio (P) e colheita (C) arroz (Fonte: CONAB, 2016).

Cultura Agrícola	23/09 a 20/12			21/12 a 20/03			20/03 a 21/06			21/06 a 23/09		
	Primavera			Verão			Outono			Inverno		
	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set
Arroz	P	P	P	C	C	C	C	C			P	P

O NDVI (*Normalized Difference Vegetation Index*) foi calculado a partir das bandas do vermelho (pRED) (620-670 nm) e infravermelho próximo (pNIR) (841-875 nm). As imagens NDVI foram agrupadas em um cubo temporal, no qual os eixos “x” e “y” representam às coordenadas geográficas e o eixo “z” às assinaturas temporais NDVI ordenadas de forma crescente.

As séries temporais MODIS são susceptíveis a ruídos provenientes de problemas na aquisição dos dados, tais como presença de nuvens, sombras e variações atmosféricas. O filtro Savitzky-Golay (SAVITZKY & GOLAY, 1964) foi aplicado para reduzir os ruídos. Este método proporciona uma suavização e eliminação dos ruídos da curva fenológica, mas busca

preservar as feições de altura do pico e largura da curva temporal (CHEN *et al.*, 2004; FONTANA *et al.*, 2008; HÜTTICH *et al.*, 2009; REN *et al.*, 2008) (Figura 4.3).

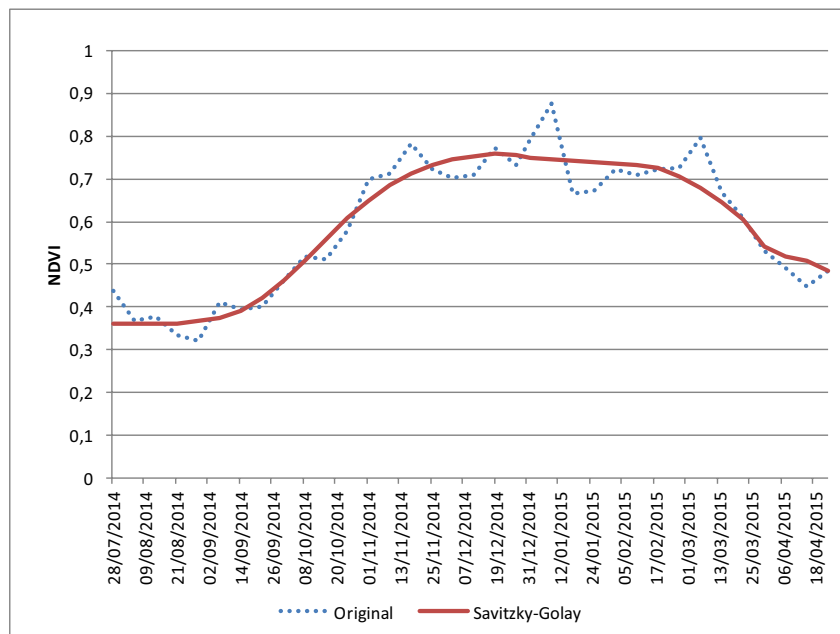


Figura 4. 3. Curva da série temporal original e com a aplicação do filtro Savitzky-Golay.

4.3.2. Classificação Baseada no Método do Vizinho Mais Próximo

A região de estudo é bastante fragmentada e possui áreas de diferentes alvos, tais como vegetações naturais, construções e plantios. Nesse contexto, foi desenvolvido um método para mapear unicamente a classe de arroz, não sendo necessário coletar amostras das diferentes classes de cobertura e uso da terra presente na região. Além disso, o método considerou a alta variação das amostras de treinamentos, pois uma mesma plantação pode possuir diferentes variações fenológicas devido aos fatores climáticos, ambientais e da própria decisão do agricultor. A metodologia foi subdividida em 3 etapas: (a) seleção de curvas fenológicas da cultura de arroz, (b) geração da imagem de distância ou similaridade baseada no método do vizinho mais próximo, e (c) definição de uma máscara de arroz a partir da definição de um limite de corte ótimo da imagem de métrica do vizinho mais próximo.

A seleção das assinaturas temporais dos plantios de arroz considerou as diferentes curvas fenológicas presentes na área de estudo. Conforme Teluguntla *et al.*, (2015), o NDVI começa a aumentar durante a fase de crescimento dos campos de arroz e atinge o auge quando muda da fase de crescimento vegetativo para crescimento reprodutivo, depois as folhas começam a murchar e finalmente com a colheita o NDVI diminui abruptamente. Portanto, os perfis temporais NDVI fornecem informações como: a intensidade dos cultivos (cultura única ou dupla), o vigor e saúde da cultura (indicada pela magnitude do NDVI) e o calendário da cultura (que mostra quando a cultura começa a emergir e quando é colhida) (GUMMA *et al.*, 2011). Cada classe de arroz tem características temporais próprias representadas pelo comportamento sazonal, início do plantio, época de colheita e pico NDVI (TELUGUNTILA *et al.*, 2015).

O procedimento de classificação desse trabalho baseou-se no método do vizinho mais próximo, um caso específico do método K-Nearest Neighbor (K-NN). O classificador K-NN é um método não paramétrico que determina o pertencimento de uma classe pelas k amostras de treinamento mais próximas do comportamento do dado analisado (COVER & HART, 1967). O método do vizinho mais próximo é quando “k” igual a 1. Essa característica permite

predizer o alvo considerando um conjunto de amostragem de treinamento com alta variação interna, como são relatadas nas séries temporais de alvos agrícolas.

O cálculo do vizinho mais próximo foi realizado a partir de duas métricas: distância Euclidiana e semelhança de cossenos. A distância Euclidiana é a métrica mais difundida e aplicada na classificação K-NN. A semelhança de cossenos é uma métrica de similaridade amplamente difundida no sensoriamento remoto pelo método Spectral Angle Mapper (KRUSE *et al.*, 1993). No entanto, essa métrica não tem sido utilizada no método K-NN. A semelhança de cossenos calcula o ângulo entre os vetores das séries temporais do pixel com a da amostra de treinamento. Portanto, imagens de métricas (distância e similaridade) foram geradas a partir do comportamento temporal de cada pixel e sua amostra de treinamento mais próxima. Os pixels da imagem resultante com maior probabilidade da presença do alvo são caracterizados por baixos valores (menor distância ou ângulo).

A determinação do melhor valor limite nas imagens de métricas entre as áreas de arroz e não-arroz utilizou a metodologia desenvolvida por Carvalho Júnior *et al.*, (2015). O método estabelece uma sequência de índices Kappa entre um mapa de referência e máscaras de arroz produzidas por uma sucessão de valores limites aplicadas na imagem de métrica. O valor limite ótimo foi determinado pela máscara de arroz com maior índice Kappa. A imagem de referência foi elaborada a partir da interpretação visual de imagens Landsat 8, com melhor resolução espacial.

4.3.3. Análise da Acurácia

A análise de acurácia da classificação foi avaliada pela matriz de confusão, precisão geral e índice Kappa (CONGALTON & GREEN, 1993). Essa avaliação foi realizada em uma área que representa 10% do total da área de estudo, entre os dados da classificação de arroz com as imagens NDVI/MODIS e o mapeamento por meio de interpretação visual das imagens do Landsat 8. Além disso, utilizamos os testes McNemar para medir a significância estatística das diferenças analisadas nas precisões das classificações. Estes testes foram elaborados para estimar diferenças entre proporções que não são independentes e tem sido bastante utilizado na detecção remota. E ainda permite comparar todas as combinações em pares de dois mapas usando o mesmo conjunto de validação (MCNAMER, 1947; FOODY, 2004; LEEUW *et al.*, 2006). No teste do McNemar, foram utilizadas 1000 amostras.

4.4. RESULTADOS

Diferentes perfis temporais NDVI de arroz foram selecionados na área de estudo, totalizando 24 amostras. A **Figura 4.4** demonstra as curvas fenológicas do arroz, onde existe uma diferenciação da data do início do plantio (aumento do NDVI) que variam de agosto a novembro e da colheita (com a queda do NDVI) que variam de fevereiro a abril. Essas diferentes curvas fenológicas podem ser explicadas pelo tipo de manejo que as lavouras recebem, além das características das lavouras, localização, datas de emergência e cultivares utilizadas (STEINMETZ *et al.*, 2015). Ainda pode ser observado que os valores de NDVI na semeadura até a emergência das plântulas ficaram em torno de 0,2 a 0,3, os valores do estágio vegetativo ficaram em torno de 0,3 a 0,4, no estágio reprodutivo os valores se encontram acima de 0,7, e os valores máximos de NDVI, que é o estágio de maturação, estão acima de 0,8. Em seguida, quando começa a colheita, ocorre o decréscimo desses valores de NDVI. Bariani *et al.* (2017) encontraram valores semelhantes para as assinaturas temporais na cultura de arroz em uma região do Rio Grande do Sul.

A **Figura 4.5** demonstra as curvas dos índices Kappa para detecção do melhor valor limite na delimitação dos plantios de arroz a partir das imagens de métricas do vizinho mais próximo. O método utilizando a distância Euclidiana obteve um valor limite ótimo em 0,488 com índice Kappa de 75,10, enquanto o método de semelhança de cossenos apresentou como

valor limite ótimo o ângulo de 0,116 com um índice Kappa de 74,06. A **Figura 4.6** demonstra os resultados da distribuição espacial da cultura do arroz ano safra 2014/2015 a partir dos limites ótimos obtidos para as duas métricas de distância Euclidiana e semelhança de cossenos. O resultado da classificação de arroz confirma a maior intensidade de cultivo na microrregião de Araranguá, responsável por 35% da produção do Estado.

A estimativa da cultura do arroz para a região mapeada foi comparada com os dados do IBGE (**Tabela 4.2**). Os resultados das classificações apresentaram valores um pouco superior do que o IBGE, 1.871 hectares a mais no método usando distância Euclidiana e 2.971 hectares pelo método de semelhança de cossenos. Esses valores representam menos do que 2% da área total de plantio, evidenciando a relevância dessa técnica para obtenção de estimativas de áreas plantadas.

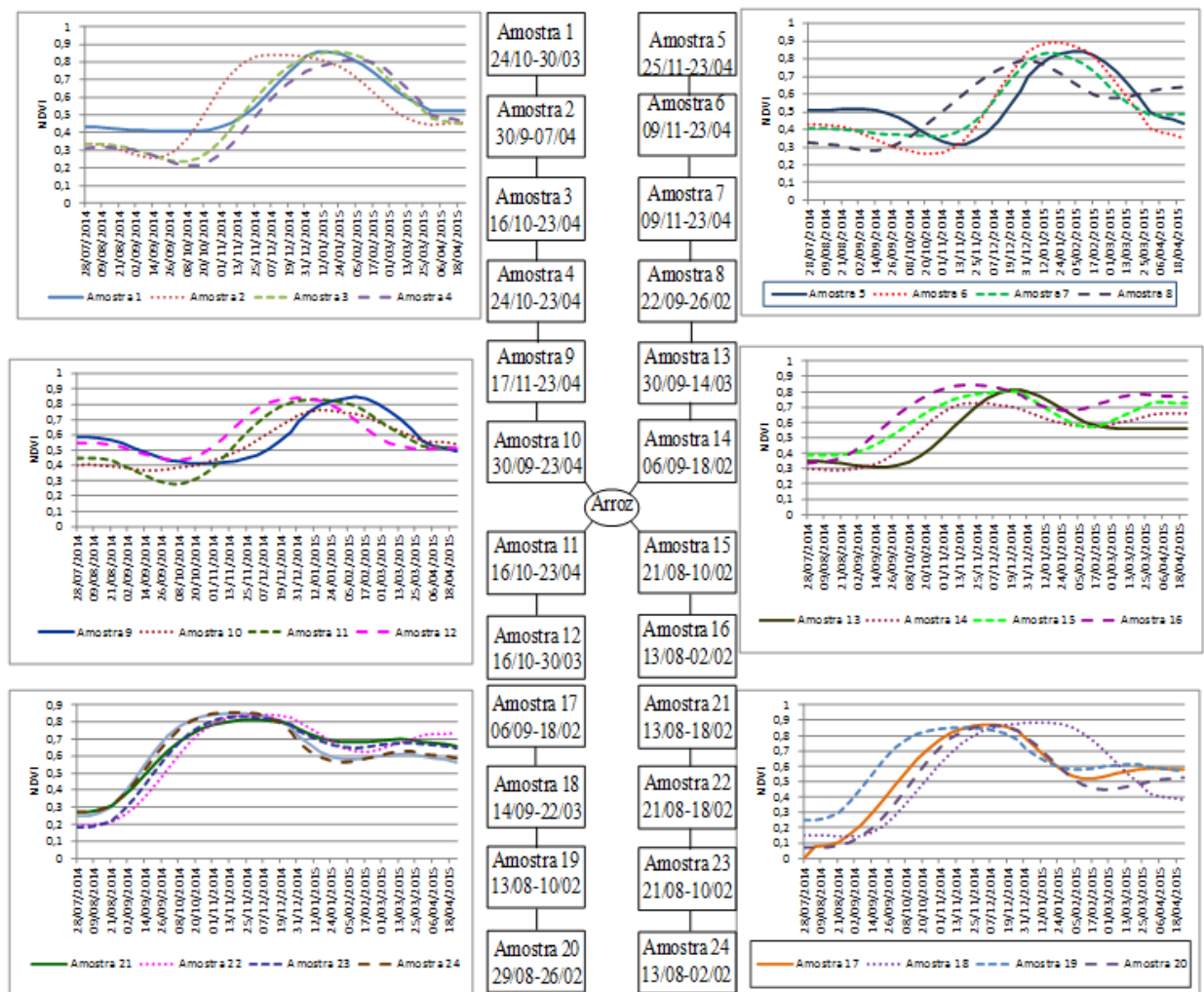


Figura 4. 4. Variação temporal dos perfis de NDVI para cada grupo de arroz.

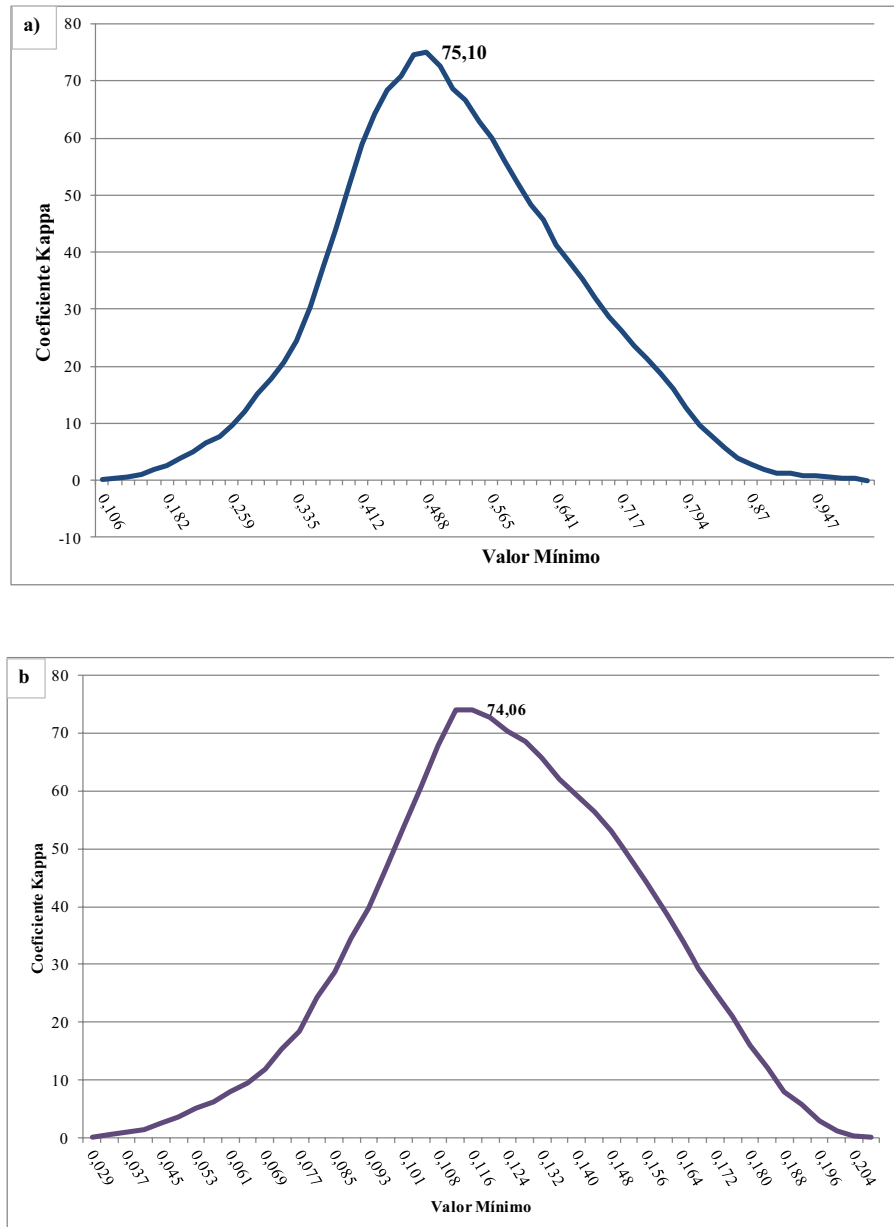


Figura 4. 5. Valor limite para a detecção da máscara de arroz a partir da imagem da métrica do vizinho mais próximo: (a) distância Euclidiana, e (b) semelhança de cossenos.

A **Tabela 4.3** apresenta a matriz de confusão das classificações em relação ao mapeamento por interpretação visual da imagem Landsat 8. A classificação baseada na distância Euclidiana obteve um índice Kappa de 0,73 e um coeficiente de precisão global de 90,82%, levemente superior ao método baseado na semelhança de cossenos com índice Kappa de 0,72 e coeficiente de precisão global de 91,21%. Os erros de omissão foram de 15,13% para o resultado da classificação com a Distância Euclidiana e de 16,58% para semelhança de cossenos, o que significa que estas áreas deveriam ter sido classificadas como arroz e não foram. No entanto, os erros de comissão para Distância Euclidiana e semelhança de cosseno foram de 25,76% e 27 % para a classe de arroz, respectivamente, ou seja, essas áreas foram incluídas erroneamente na classificação.

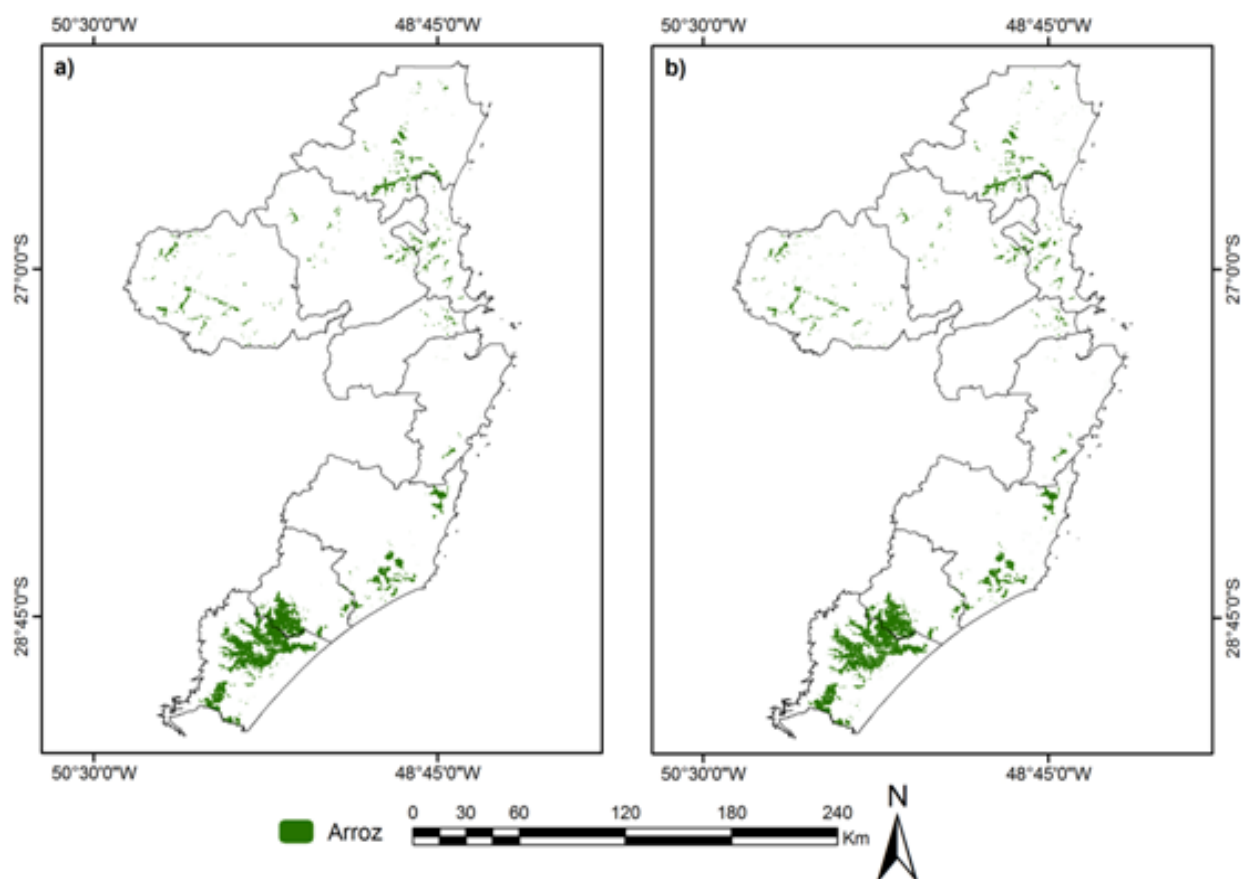


Figura 4. 6. Área de arroz classificada com: a) Distância Euclidiana e b) semelhança de cossenos.

Tabela 4. 2. Área cultivada com arroz em hectares, estimada com as duas metodologias baseada nas métricas do vizinho mais próximo: distância euclidiana (DE) e semelhança de cossenos (SC).

Área de estudo	IBGE (2013)	DE	SC
Área Total	147.775	149.646	150.746

Tabela 4. 3. Matriz de confusão da classificação baseada nas métricas de distância e similaridade do vizinho mais próximo (%).

Distância Euclidiana				Semelhança de Cossenos			
Overall Accuracy		91,21		Overall Accuracy		90,82	
Kappa Coefficient		0,73		Kappa Coefficient		0,72	
Classes	Arroz	Não-Arroz	Total	Classes	Arroz	Não-Arroz	Total
Arroz	84,87	7,23	22,53	Arroz	83,42	7,40	22,11
Não-Arroz	15,13	92,77	77,47	Não-Arroz	16,58	92,60	77,89
Total	100	100	100	Total	100	100	100
Erro de Omissão	15,13	7,23	-	Erro de Omissão	16,58	7,40	-
Erro de Comissão	25,76	3,85	-	Erro de Comissão	27,00	4,12	-

As duas classificações pelo vizinho mais próximo (distância Euclidiana e similaridade do cosseno) tiveram igualdade estatística dos resultados como certifica o teste de McNemar, contendo χ^2 de 1.689 menor que o valor crítico de qui-quadrado de 3,841 referente a 5% de probabilidade (**Tabela 4.4**). Além disso, este teste baseou-se em matriz de confusão de duas dimensões, onde foi focada a distinção binária entre a alocação das classes correta e incorreta. A tabela demonstrou que 97,10 % dos pares eram iguais, correto x correto e incorreto x incorreto.

Tabela 4. 4. Teste de McNemar's entre as duas classificações do vizinho mais próximo usando as métricas de distância Euclidiana e semelhança de cossenos

	Distância Euclidiana		
	Correto	Incorreto	Total
Semelhança de Cossenos			
Correto	897	18	915
Incorreto	11	74	85
Total	908	92	1000
χ^2	1.689		

4.5. DISCUSSÃO

As formas das séries temporais NDVI são utilizadas para diferenciar as fases de crescimento de diversos cultivares (GUMMA *et al.*, 2011; WANG *et al.*, 2014; BARIANI *et al.*, 2017). As variações dos perfis fenológicos de arroz presentes na área de estudo foram constatadas em outros estudos. Gumma *et al.*, (2014) demonstraram que o uso do MODIS captura a variabilidade sazonal na extensão e área da cultura do arroz. Teluguntla *et al.*, (2015) constataram que o estudo da fenologia permite a separação do arroz de curto e longo período com base em fatores como início da safra e a magnitude do NDVI. Estel *et al.*, (2016) demonstraram diversos sistemas de cultivos de arroz utilizando indicadores como frequência de cultivo, cultivo múltiplo, pousio dos ciclos e taxa de duração da cultura. Todas essas variações devem ser consideradas na metodologia de classificação usando perfis fenológicos (GUMMA *et al.*, 2011; PENG *et al.*, 2011; SINGHA *et al.*, 2016). A diferença do comportamento temporal do arroz está relacionada a diferentes fatores, tais como a localização geográfica, a cultivar utilizada (devido ao ciclo biológico de cada variedade que influencia no desenvolvimento da planta), o clima (escassez de chuva ou irregularidade na distribuição) e prática de manejo (STEINMETZ *et al.*, 2015).

A classificação baseada na métrica do método K-NN (distancia ou similaridade), aplicada neste trabalho, considerou amostras de treinamento com assinaturas diversas da cultura de arroz. Este procedimento tornou o método adaptável e apropriado para estudos de alvos que descrevem variações espaços temporais. Geralmente as áreas de estudo são bastante fragmentadas e com diversos alvos, no caso do método K-NN, apenas um único alvo foi classificado, pois foi possível adquirir valores similares das assinaturas temporais distintas, considerando diferentes datas de plantio.

Mapear áreas de cultivos de arroz é de suma importância devido ao grande consumo do grão no país. O mapeamento da cultura do arroz, na região de Santa Catarina, mostrou valores de precisão equivalentes com outros estudos com sensor MODIS e a área plantada alcançou valores estimados por estatísticas de órgãos governamentais, como os encontrados nos trabalhos de Gumma *et al.*, (2011) e Teluguntla *et al.*, (2015). Os resultados provenientes da classificação das imagens MODIS NDVI, compõem erros devido a fatores relacionados a pixels mistos causados pela resolução grosseira destes e também por causa da extensão das

propriedades rurais que são de tamanho médio de 13,5 ha. Entretanto, o uso de sensores com alta resolução espacial e temporal podem diminuir esses erros de classificação.

4.6. CONCLUSÃO

O estudo detectou a distribuição espacial do plantio de arroz no Estado de Santa Catarina a partir de séries temporais MODIS NDVI. O método de detecção proposto foi baseado nas métricas do vizinho mais próximo e o estabelecimento de uma máscara do plantio de arroz a partir de um valor limite. Esse método demonstrou adequado para alvos que apresentam significativas variações espaço-temporais e presença de outros alvos desconhecidos. A composição das imagens de 8 dias permitiu captar com precisão todo o ciclo vegetativo da cultura de arroz, desde plantio a colheita. Os índices Kappa das classificações foram muito próximos, sendo 0,73 para a distância Euclidiana e 0,72 para a semelhança de cossenos. O teste de McNemar demonstrou que os resultados das duas classificações são equivalentes. As estimativas da área de arroz pelos dois métodos apresentaram resultados similares aos dados oficiais do governo. Portanto, o método auxilia o sistema de previsão de safras do governo e monitoramento das áreas de arroz. Dados estatísticos acurados e recentes são elementos primordiais para o estabelecimento de estratégias de desenvolvimento econômico e de segurança alimentar.

AGRADECIMENTOS

A CONAB pelo apoio e por fornecer os pontos georreferenciados das lavouras de arroz.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARIANI, C. J. M. V.; BARIANI, N. M. V.; MARQUES NETO, G. C. M. Sensoriamento remoto aplicado ao monitoramento dos principais estádios fenológicos do arroz irrigado no Sul do Brasil. In: **IV Inovagri International Meeting**, 2017, Fortaleza. Anais do IV Inovagri International Meeting - 2017.
- CARVALHO JÚNIOR, O. A. de; GUIMARÃES, R. F.; SILVA, C. R.; GOMES, R. A. T. Standardized Time-Series and Interannual Phenological Deviation: New Techniques for Burned-Area Detection Using Long-Term MODIS-NBR Dataset. **Remote Sensing**, v. 7, n. 6, p. 6950–6985, 2015.
- CHEN, C. F.; SON, N. T.; CHANG, L. Y. Monitoring of rice cropping intensity in the upper Mekong Delta, Vietnam using time-series MODIS data. **Advances in Space Research**, v. 49, n. 2, p. 292–301, 2012.
- CHEN, J.; JÖNSSON, P.; TAMURA, M.; GU, Z.; MATSUSHITA, B.; EKLUNDH, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. **Remote Sensing of Environment**, v. 91, n. 3–4, p. 332–344, 2004.
- CLAUSS, K.; YAN, H.; KUENZER, C. Mapping paddy rice in China in 2002, 2005, 2010 and 2014 with MODIS time series. **Remote Sensing**, v. 8, n. 5, p. 1–22, 2016.
- COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). **A cultura do arroz**. Brasília: Conab, 2015.
- COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). **Acompanhamento da**

safr **brasileira** **de** **grãos**, 2016. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/16_06_13_12_32_09_boletim_gaos_janeiro_2016.pdf>. Acesso em: 30 jan. 2017.

CONGALTON, R. G.; GREEN, K. A practical look at the sources of confusion in error matrix generation. **Photogrammetric Engineering & Remote Sensing**, v. 59, n. 5, p. 641–644, 1993.

COVER, T.; HART, P. Nearest neighbor pattern classification. **IEEE Transactions on Information Theory**, v. 13, n. 1, p. 21–27, 1967. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1053964>>.

D'ARCO, E.; ALVARENGA, B. S.; RIZZI, R.; FRIEDRICH, B.; RUDORFF, T.; MOREIRA, M. A.; ADAMI, M. Geotecnologias na estimativa da área plantada com arroz irrigado. **Revista Brasileira de Cartografia**, v. 58, n. 3, p. 247–253, 2006.

EMPRESA DE PESQUISA AGROPECUÁRIA E EXTENSÃO RURAL DE SANTA CATARINA (EPAGRI). **Síntese Anual da Agricultura de Santa Catarina 2015 - 2016**. Disponível em: <http://docweb.epagri.sc.gov.br/website_cepa/publicacoes/Sintese_2016.pdf>. Acesso em: 30 jan. 2017.

FONTANA, F.; RIXEN, C.; JONAS, T.; ABEREGG, G.; WUNDERLE, S. Alpine grassland phenology as seen in AVHRR, VEGETATION, and MODIS NDVI time series - A comparison with in situ measurements. **Sensors**, v. 8, n. 4, p. 2833–2853, 2008.

FOODY, G. M. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. **Photogrammetric Engineering & Remote Sensing**, v. 70, n. 5, p. 627–633, 2004.

GUAN, X.; HUANG, C.; LIU, G.; MENG, X.; LIU, Q. Mapping rice cropping systems in Vietnam using an NDVI-based time-series similarity measurement based on DTW distance. **Remote Sensing**, v. 8, n. 1, 2016.

GUINDIN-GARCIA, N.; GITELSON, A. A.; ARKEBAUER, T. J.; SHANAHAN, J.; WEISS, A. An evaluation of MODIS 8- and 16-day composite products for monitoring maize green leaf area index. **Agricultural and Forest Meteorology**, v. 161, p. 15–25, 2012.

GUMMA, M. K.; MOHANTY, S.; NELSON, A.; ARNEL, R.; MOHAMMED, I. A.; DAS, S. R. Remote sensing based change analysis of rice environments in Odisha, India. **Journal of Environmental Management**, v. 148, p. 31–41, 2015.

GUMMA, M. K.; NELSON, A.; THENKABAIL, P. S.; SINGH, A. N. Mapping rice areas of South Asia using MODIS multitemporal data. **Journal of Applied Remote Sensing**, v. 5, n. 1, p. 53547, 2011.

GUMMA, M. K.; THENKABAIL, P. S.; MAUNAHAN, A.; ISLAM, S.; NELSON, A. Mapping seasonal rice cropland extent and area in the high cropping intensity environment of Bangladesh using MODIS 500m data for the year 2010. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 91, p. 98–113, 2014.

HÜTTICH, C.; GESSNER, U.; HEROLD, M.; STROHBACH, B. J.; SCHMIDT, M.; KEIL, M.; DECH, S. On the suitability of MODIS time series metrics to map vegetation types in dry savanna ecosystems: A case study in the Kalahari of NE Namibia. **Remote Sensing**, v. 1, n. 4, p. 620–643, 2009.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Levantamento**

sistemático da produção agrícola. v. 26 n.9 ed. [s.l: s.n.], 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Levantamento sistemático da produção agrícola** Rio de Janeiro, 2016. Disponível em: <[ftp://ftp.ibge.gov.br/Producao_Agricola/Levantamento_Sistematico_da_Producao_Agricola_\[mensal\]/Fasciculo/lspa_201612.pdf](ftp://ftp.ibge.gov.br/Producao_Agricola/Levantamento_Sistematico_da_Producao_Agricola_[mensal]/Fasciculo/lspa_201612.pdf)>. Acesso em: 27 jan. 2017.

JIN, C.; XIAO, X.; DONG, J.; QIN, Y.; WANG, Z. Mapping paddy rice distribution using multi-temporal Landsat imagery in the Sanjiang Plain, northeast China. **Frontiers of Earth Science**, v. 10, n. 1, p. 49–62, 2015.

KLERING, E. V.; FONTANA, D. C.; ALVES, R.; ROCHA, J.; BERLATO, M. A. Estimativa de Área Cultivada com Arroz Irrigado para o Estado do Rio Grande do Sul a partir de Imagens Modis. **Ciência e Natura**, v. 35, n. 2, p. 126–135, 2014.

KONTGIS, C.; SCHNEIDER, A.; OZDOGAN, M. Mapping rice paddy extent and intensification in the Vietnamese Mekong River Delta with dense time stacks of Landsat data. **Remote Sensing of Environment**, v. 169, p. 255–269, 2015.

KRUSE, F. A.; LEFKOFF, A. B.; BOARDMAN, J. W.; HEIDEBRECHT, K. B.; SHAPIRO, A. T.; BARLOON, P. J.; GOETZ, A. F. H. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. **Remote Sensing of Environment**, v. 44, n. 2–3, p. 145–163, 1993.

LEEuw, J. D.; JIA, H.; YANG, L.; SCHMIDT, K.; SKIDMORE, A. K. Comparing accuracy assessment to infer superiority of image classification methods. **International Journal of Remote Sensing**, v. 27, n. 1, p. 223–232, 2006.

MACLEAN, J.; DAWE, D.; HARDY, B.; HETTEL, G. **Rice Almanac: Source book for the most important economic activity on earth** CABI Publishing Wallingford, UK, 2002.

MCNEMAR, Q. Note on the sampling error of the difference between correlated proportions or percentages. **Psychometrika**, v. 12, n. 2, p. 153–157, 1947.

PANDOLFO, C.; BRAGA, H. J.; SILVA JÚNIOR, V.P. MASSIGNAN, A. M.; PEREIRA, E. S.; THOMÉ, V. M. R.; VALCI, F. V. **Atlas climatológico do Estado de Santa Catarina**. Florianópolis: Epagri, 2002. Disponível em: <http://www.ciram.sc.gov.br/atlas_climatologico/>.

PENG, D.; HUETE, A. R.; HUANG, J.; WANG, F.; SUN, H. Detection and estimation of mixed paddy rice cropping patterns with MODIS data. **International Journal of Applied Earth Observation and Geoinformation**, v. 13, n. 1, p. 13–23, 2011.

POTTER, R. O.; CARVALHO, A. P. de; FLORES, C. A.; BOGNOLA, I. Solos do estado de Santa Catarina. **Boletim de Pesquisa e Desenvolvimento**, v. 46, 2004.

REN, J.; CHEN, Z.; ZHOU, Q.; TANG, H. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. **International Journal of Applied Earth Observation and Geoinformation**, v. 10, n. 4, p. 403–413, 2008.

RIZZI, R.; RUDORFF, B. F. T. Estimativa da área de soja no Rio Grande do Sul por meio de imagens landsat. **Revista Brasileira de Cartografia**, v. 57, n. 3, p. 226–234, 2005.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627–1639, 1964.

SINGHA, M.; WU, B.; ZHANG, M. An object-based paddy rice classification using multi-spectral data and crop phenology in Assam, Northeast India. **Remote Sensing**, v. 8, n. 6,

2016.

SOCIEDADE SUL-BRASILEIRA DE ARROZ IRRIGADO (SOSBAI). Arroz Irrigado : Recomendações Técnicas da Pesquisa para o Sul do Brasil / XXX Reunião Técnica da Cultura do Arroz Irrigado, 06 a 08 de agosto de 2014, Bento Gonçalves, RS, Brasil. p. 192, 2014.

SUBUDHI, P. K.; SASAKI, T.; KHUSH, G. S. 1 Rice. **Genome Mapping and Molecular Breeding in Plants, Volume 1 Cereals and Millets**, v. 1, 2006.

STEINMETZ, S.; TEMPERADO, C.; CUADRA, S. V.; TEMPERADO, E. C.; ALMEIDA, I. R. De. GD Arroz: Programa Baseado em Graus-Dia como Suporte ao Planejamento e à Tomada de Decisão no Manejo do Arroz Irrigado. 2015.

TELUGUNTLA, P.; RYU, D.; GEORGE, B.; WALKER, J.; MALANO, H. Mapping Flooded Rice Paddies Using Time Series of MODIS Imagery in the Krishna River Basin, India. **Remote Sensing**, v. 7, n. 7, p. 8858–8882, 2015.

WANG, L.; ZHANG, F. C.; JING, Y. S.; JIANG, X. D.; YANG, S. Bin; HAN, X. M. Multi-temporal detection of rice phenological stages using canopy stagespectrum. **Rice Science**, v. 21, n. 2, p. 108–115, 2014.

CAPÍTULO 5– CONSIDERAÇÕES FINAIS

O agronegócio brasileiro ganhou participação significativa no mercado internacional de alimentos, graças ao processo de transformação da agricultura brasileira que possibilitou a modernização e expansão da agricultura. Como a agricultura é importante para a economia do país, e devido a sua dinâmica espaço-temporal, torna-se difícil a aquisição de informações confiáveis sobre este setor. Por outro lado, o uso do sensoriamento remoto se mostra eficaz e contribui na obtenção de informações de modo rápido, preciso e com baixo custo.

Deste modo, os dados multiespectrais e multitemporais do sensor MODIS são fundamentais para o mapeamento e monitoramento de culturas agrícolas em grandes regiões produtoras como no Brasil. Assim, o empenho da geotecnologia aplicada em mapeamentos de diversas culturas agrícolas no Brasil e no mundo contribui para a ciência geográfica no que se refere ao uso e ocupação do solo ao longo do tempo e do espaço. Portanto, com o uso do produto MOD09Q1 do sensor MODIS de 8 dias e com resolução de 250m, foi possível identificar e mapear as diversas culturas agrícolas nas três áreas de estudo (Capítulo 2, 3 e 4).

Para alcançar os objetivos desta tese, buscou-se comparar métodos de classificações nas séries temporais do produto MOD09Q1, que poderá ser usada como auxílio para acompanhamento do crescimento das lavouras e monitoramento de previsões de safras das principais culturas agrícolas do país.

Como a série temporal MODIS é afetada por uma quantidade de ruídos, foi utilizado o filtro Savitzky-Golay que gerou resultados satisfatórios com a suavização das curvas dos espectros temporais. Estes resultados podem ser vistos nos capítulos 2, 3 e 4.

No processo de análise dos dados utilizados nesta tese, a avaliação da precisão mostrou-se confiável. Para os resultados das classificações, a matriz de confusão e o índice Kappa forneceram dados que descreveram a precisão e caracterizaram os erros, podendo aprimorar a classificação (Capítulo 2, 3 e 4). Esses resultados foram confirmados pelo teste McNemar.

Os resultados dessa tese mostraram que as séries temporais NDVI MODIS possuem uma grande capacidade para diferenciar os padrões fenológicos de cada cobertura vegetal (natural e antrópica). No Estado do Mato Grosso (Capítulo 2), existem grandes extensões de áreas agrícolas e diferentes culturas como soja, milho e algodão e grandes extensões de áreas com cobertura vegetal natural, e, foi possível identificar e separar as assinaturas temporais desses alvos, mesmo sendo complexa a discriminação das culturas agrícolas. As assinaturas temporais permitiram identificar a dinâmica sazonal das classes: formação florestal, cerrado,

pastagem, sistema anual único de cultivo (soja, milho e algodão), sistema de cultivo duplo (soja/milho e soja/algodão) e pivô central (sistema triplo de cultivo). Além da análise das assinaturas temporais, avaliamos o desempenho de dois métodos de classificação, redes neurais de retro-propagação e SVM, para mapear as grandes extensões de áreas agrícolas, e os dois métodos se mostraram apropriados para identificar as culturas agrícolas, mesmo considerando a baixa resolução espacial do MODIS.

Diferentemente da Região Centro-Oeste, as lavouras da Região Sul se destacam como culturas de pequenas propriedades rurais, com mão de obra familiar. Como estas áreas são muito fragmentadas, optou-se em desenvolver uma técnica para mapear uma única cultura agrícola (Capítulo 3 e 4), aplicando a classificação de séries temporais NDVI MODIS pelo método do vizinho mais próximo (K-NN), comparando duas métricas: similaridade e distância. Assim, foi coletada uma grande quantidade amostras de treinamento da cultura do trigo e do arroz, em razão das plantações possuírem diferentes variações fenológicas por causa das condições climáticas, ambientais, manejo da cultura e pela diversidade de cultivares. Desta forma, no Estado do Rio Grande do Sul (Capítulo 3), as assinaturas temporais da cultura do trigo possibilitaram distinguir os locais das lavouras de trigo. E no Estado de Santa Catarina (Capítulo 4), as assinaturas temporais da cultura do arroz proporcionaram localizar as lavouras de arroz. Esta técnica produziu um método flexível e ajustável para pesquisas de alvos com variações das assinaturas temporais.

A aplicação das metodologias apresentadas nesta tese obteve resultados satisfatórios, sendo possível estimar a produtividade agrícola (Capítulo 3 e 4) e detectar as diferenças nos ciclos fenológicos que estão relacionados ao calendário agrícola de cada região e cultura (Capítulo 2, 3 e 4). Desta maneira, recomenda-se que sejam aplicados em outras regiões produtoras, com outras culturas agrícolas e em outros anos safras. Entretanto, torna cada vez mais relevante o desenvolvimento de tecnologias capazes de refinar as observações ao longo de cada cultura agrícola. Portanto, propõem-se a comparação do desempenho do NDVI com outros índices, e até mesmo o desenvolvimento de novos algoritmos de classificação capazes de detectar de forma mais precisa as culturas agrícolas por intermédio do sensor MODIS.